

経情協46-13

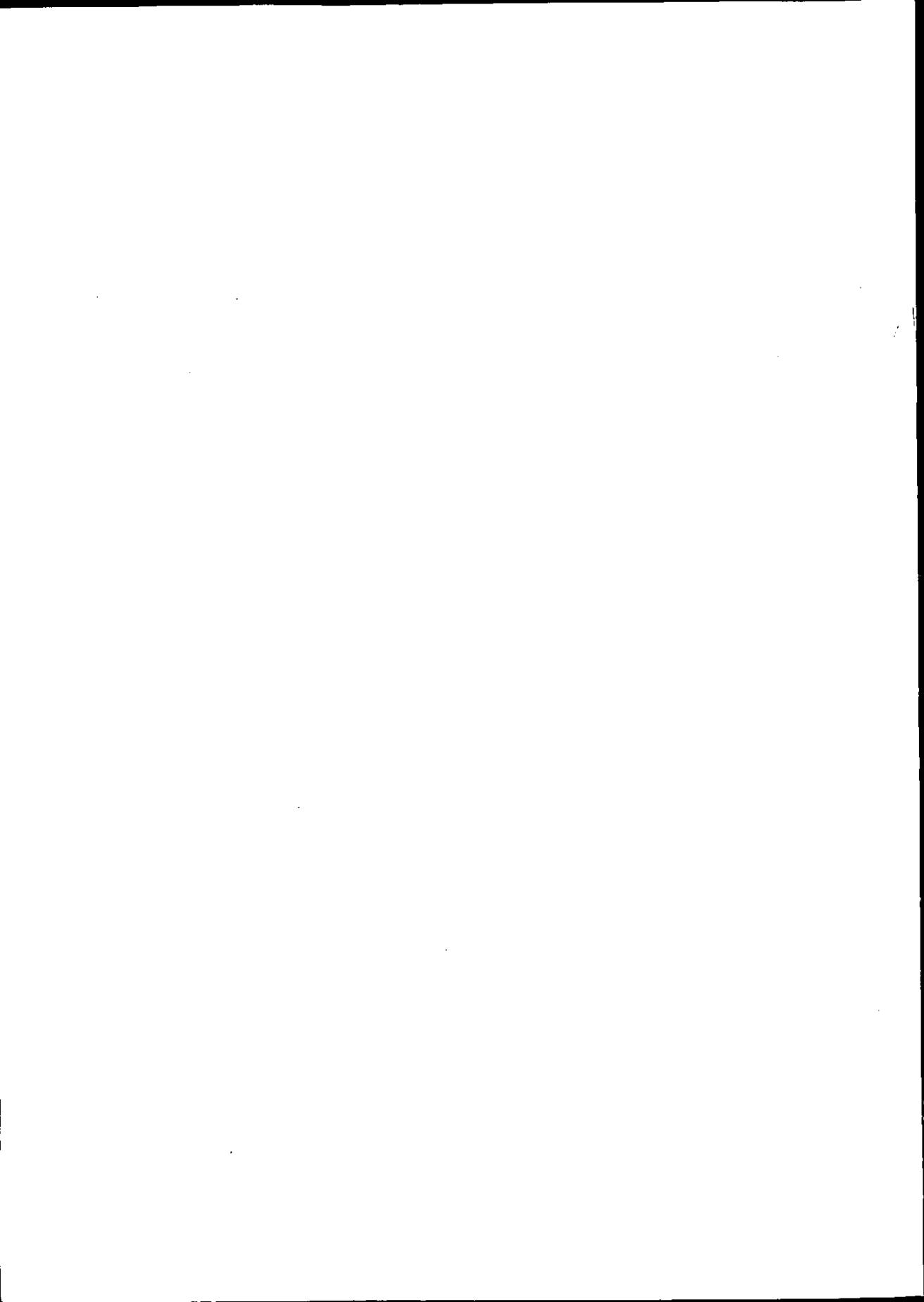
# データバンク研究報告

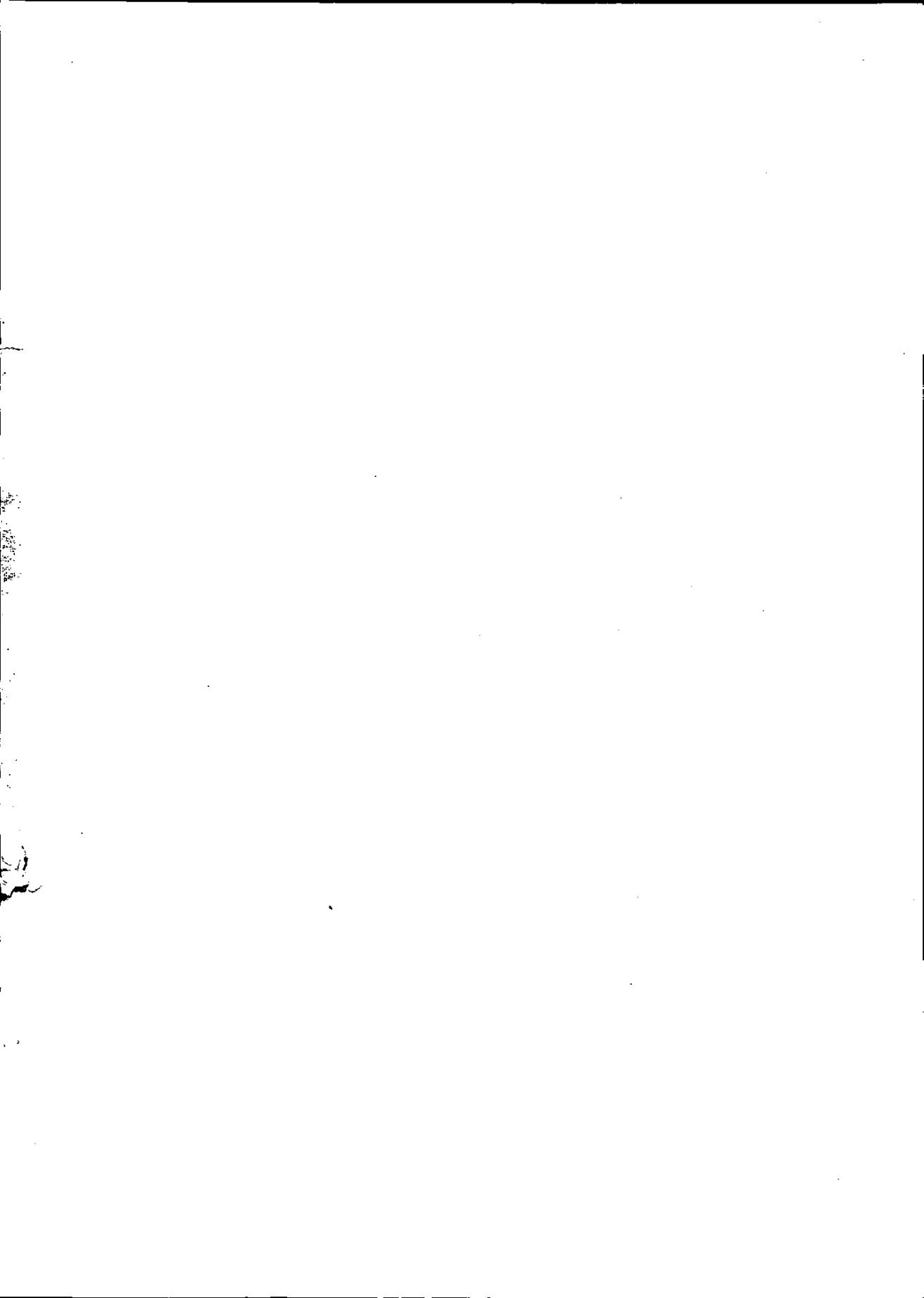
データバンク研究資料 No. 6

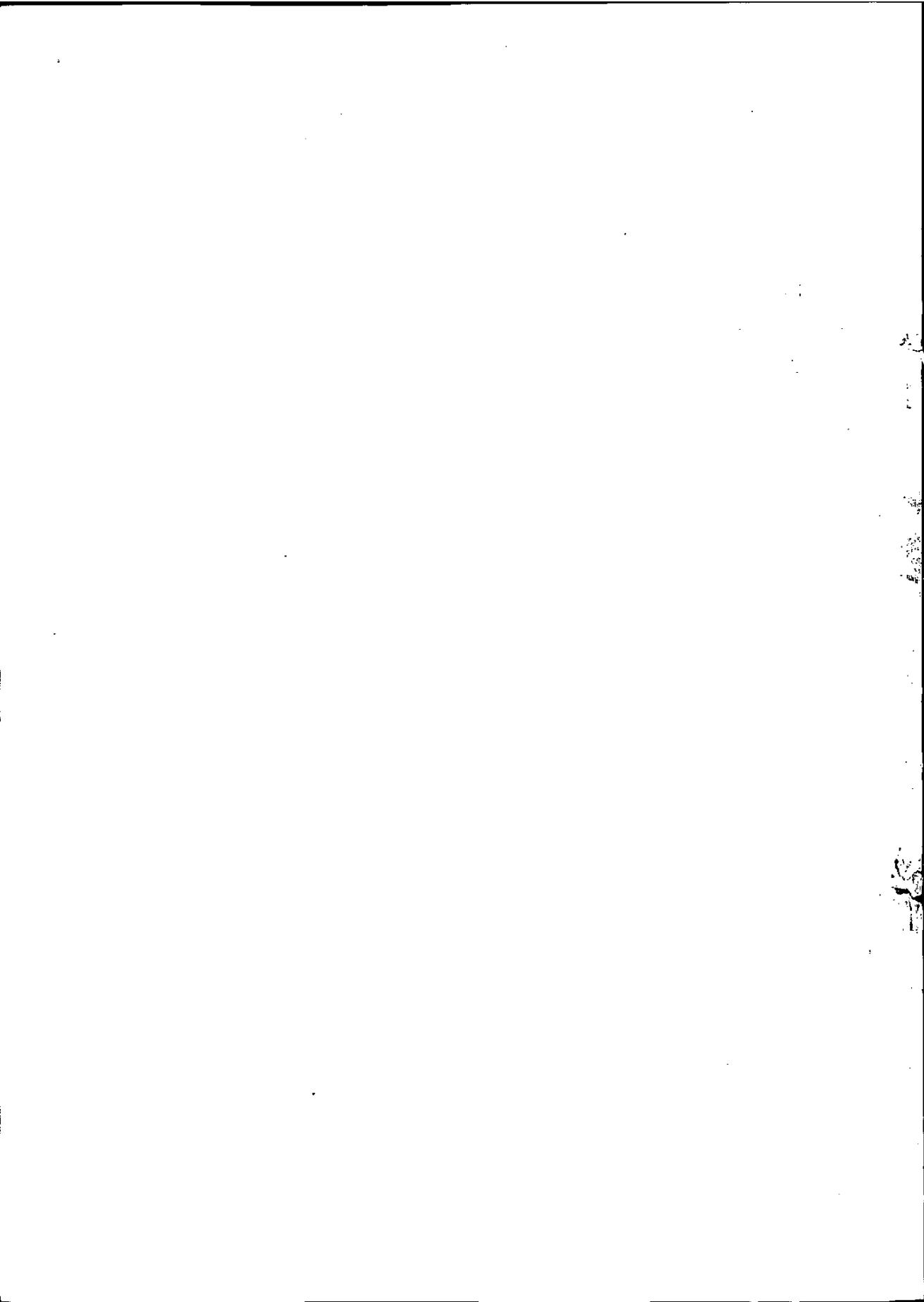
昭和46年6月

財団法人 日本経営情報開発協会









## は し が き

データバンク研究委員会が、財団法人日本経営情報開発協会に設置されたのは、1969年8月のことである。それ以来広く関係内外各方面の協力を得て1971年6月末現在まで1年10月にわたり、鋭意研究調査を継続してきた。

研究調査は、理論及び実証の両面にわたり、また国内のみならず、海外に及び、特にアメリカには1970年3月海外調査団を派遣し、さらにその結果に鑑み、内外注目のまどであるカルフォルニア州のデータ・バンクを精しく調査することとし、同州財務長官特別補佐官M・F・マクドナルド氏を煩わして1970年10月に特別の調査を委託した。

当研究委員会の調査研究の経過はすでに下記中間報告に順次発表している。

- (1) データ・バンク・シンポジウム(1970年2月)
- (2) アメリカのデータ・バンク  
—情報ネットワーク海外調査団報告書(1970年6月)
- (3) データ・バンク研究に関する中間報告書(1970年6月)
- (4) アメリカのデータ・バンクに関する資料(1971年1月)
- (5) カルフォルニア州のデータ・バンク (1971年6月)

今回の報告は、当委員会の最終報告書であり、内容は三部より構成されている。

第1部データ・バンク理論では、まずデータ・バンクをめぐる諸概念の確立・把握から出発しデータ論、ファイル構成と探索、データ処理技術、ネットワーク等、一応データ・バンクにおいてとりあぐべき諸問題をすべて、検討してある。

第2部データ・バンクの事例研究は、わが国及びアメリカのデータ・バンクの代表的な事例について、くわしく検討した。上記カルフォルニア州のハブリック・セクターにおけるデータ・ベースのほか、組織内データ・ベース、企業性及び公共性データ・バンクにわけて考察を行った。

第3部は、第1部の理論的視座において、第2部の事例研究を媒介として、データ・バンクの現状を分析評価するとともに、今後の進むべき道についての論議を展開したものである。

データ・バンクはその重要性が情報化社会の基盤として広く認識されているにもかかわらず、未だ解明すべき多くの問題が残されている。本報告書は2年近くにわたるわれわれの努力の結果ではあるが、この分野に立ち向うのには、まだ時間不足の憾みがあり、上記各部門の研究には、いまだ論究し確明すべきことが、多く残されている。しかしデータ・バンクに対する、わが国における草分けの研究として、何らかのお役に立ちうるならば、われわれの幸いとするとところである。

この報告書の内容は、精粗一様でない点もあり、まず大綱を理解していただくためには、第1部の基礎概念とデータ論を読まれ、直ちに第2部の事例研究に入られ、第3部の分析へと進まれ最後に第1部の技術論に立ち帰って御検討をいただくのがよいかと思う。

なおデータ・バンク研究委員会は下記のメンバーをもって構成されたものである。

- 北川 敏 男 (九州大学・教授)
- 足立 哲 朗 (日本興業銀行・調査部調査役)
- 今井 賢 一 (一ツ橋大学・助教授)
- 印東 太郎 (慶応義塾大学・教助)
- 河合 三 良 (行政管理庁・行政管理局長)
- 佐久間 孝 (電力中央研究所・電子計算室総括課長)
- 鈴木 康 (日本開発銀行・設備投資研究所主任研究員)
- 高瀬 保 (京都産業大学・教授)
- 中井 浩 (科学技術情報センター・資料部主任情報員)
- 野口 照 雄 (興和石油・専務取締役)
- 長谷川 寿 彦 (日本電信電話公社・データ通信本部・調査役)
- 堀 比呂志 (関西電力・企画部付)
- 渡辺 竜 雄 (通産省・情報管理課総括班長)
- 馬 越 善 通 (日本経営情報開発協会・理事)

この報告書は委員の協力により、再三の討議を重ねた結果である。特に第1部は中井浩委員の寄与によるところが多い。また本報告書の作成に当ってはジスト(日本総合技術研究所)の金成洋治氏の御援助を得たことを附記したい。

なお、この報告書は日本自転車振興会の機械工業振興資金による“経済統計データ・バンクに関する基礎調査の昭和45年度報終告”に該当するものである。

昭和46年6月

データ・バンク研究委員会

委員長 北川 敏 男

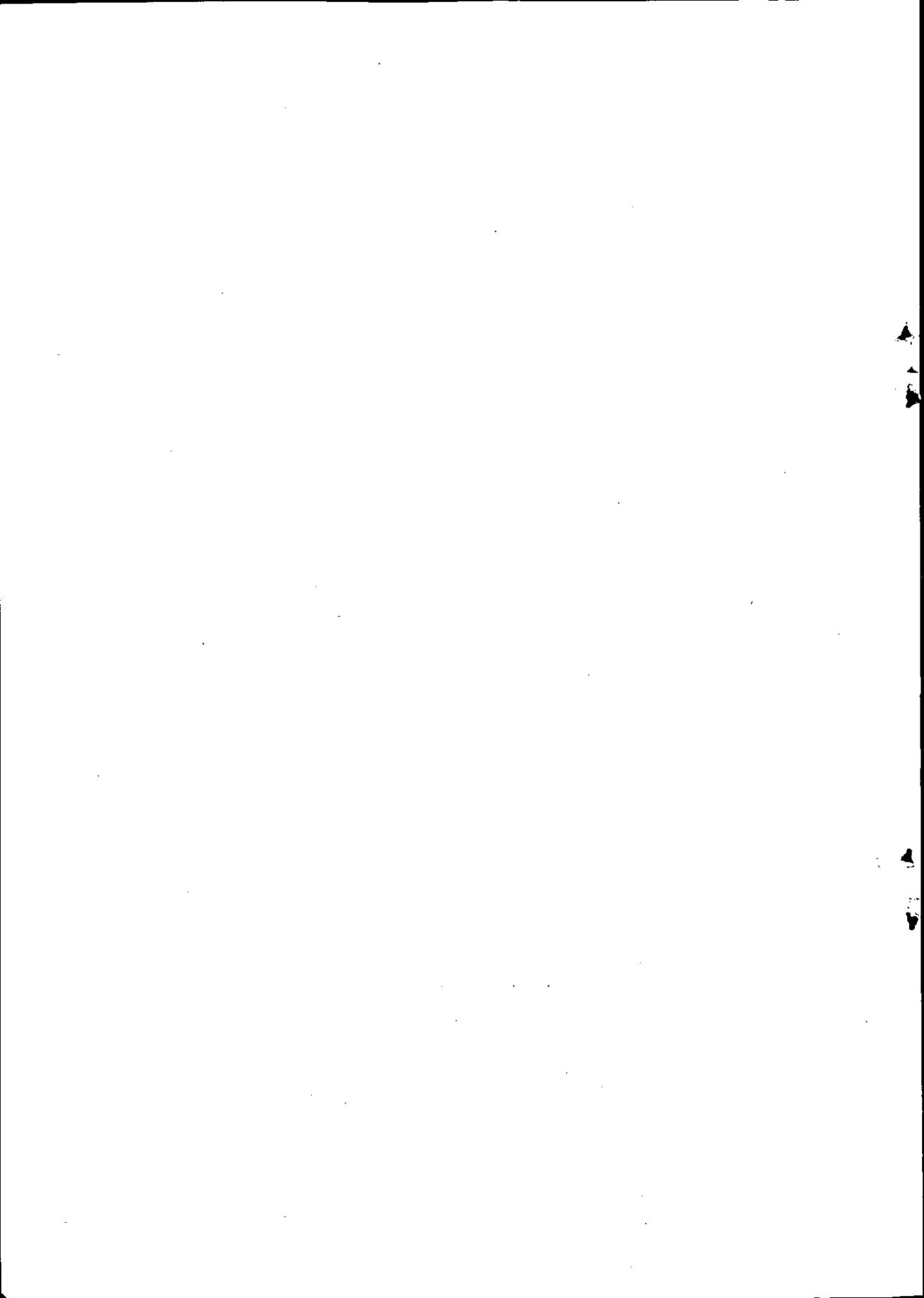
# 目 次

## 第1部 データバンクの理論

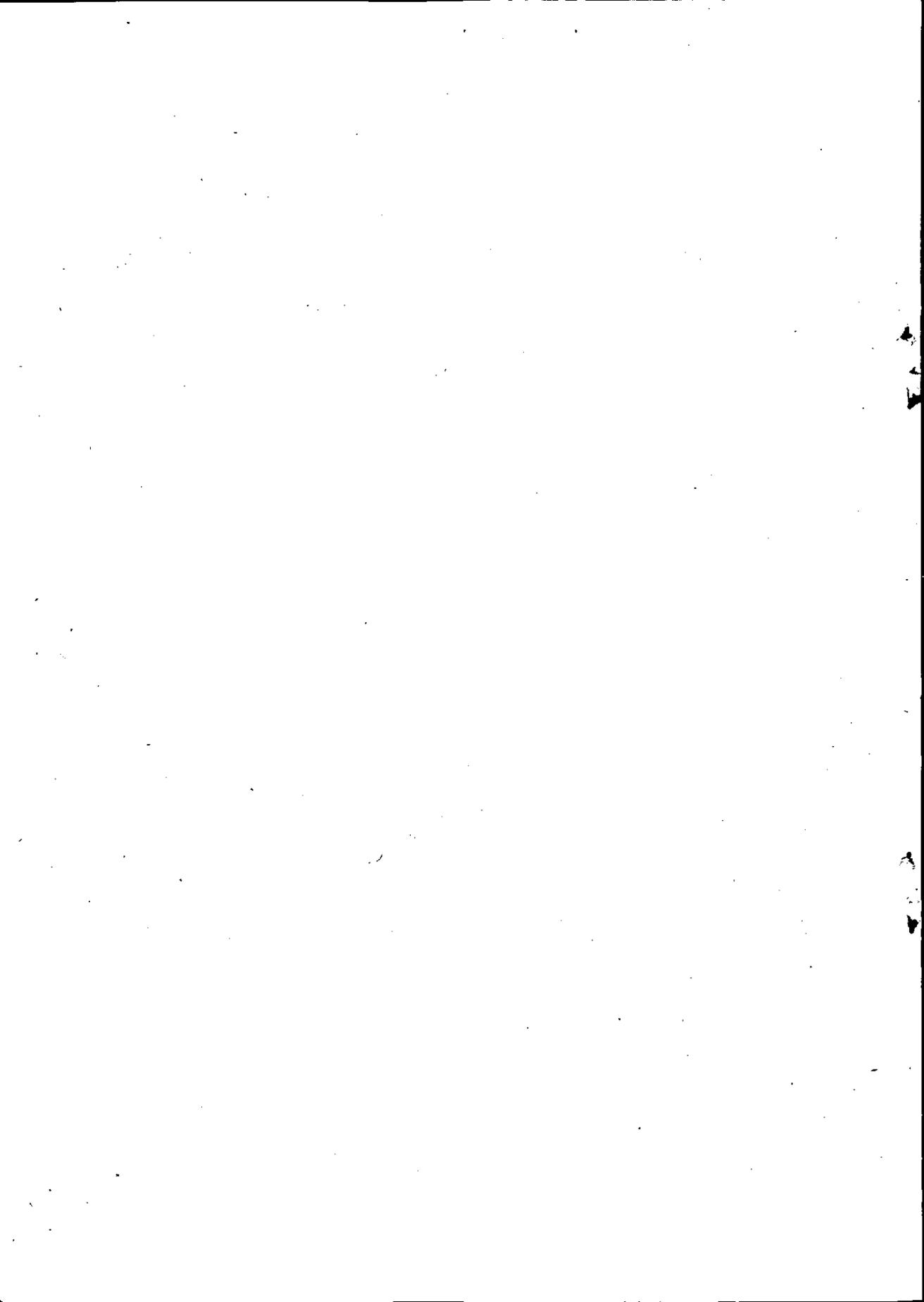
I データバンクの概念 .....	1頁
II データに対する要求の変化とデータ発生源の現状 .....	4
(1) データのもつ特性 .....	4
(2) データ需要の変化 .....	5
(3) 公共財データの発生と流通の現状と問題点 .....	8
(i) 発生と流通の現状例 .....	8
(ii) 公共財情報の問題点 .....	8
(iii) 処理手法とデータ供給 .....	14
(iv) 情報処理システムに要求される特性 .....	14
III データの収集と分析 .....	17
(1) 収 集 .....	19
(i) ファイル構成において留意すべき点 .....	20
(ii) データ源において留意すべき点 .....	20
(iii) 公共機関の責任 .....	21
(2) 分 析 .....	22
IV ファイルの構成と探索 .....	25
(1) ファイル構成の技術 .....	26
(i) データ交換用ファイル .....	26
(ii) 探索用ファイル .....	26
(2) 探 索 .....	28
(i) ファイル探索の技術 .....	28
A サ ー チ .....	28
B 照 合 .....	29
C 構造適合性の判断 .....	30
(ii) 質 問 編 集 .....	32
(iii) 回 答 編 集 .....	34
(3) プログラム言語 .....	35
(4) オンライン・システムおよびマルチ・ベースの管理問題 .....	36

i	オンライン・システムの構成	36頁
ii	会話の形式	38
V	センテシアル・データおよびアナログ・データの処理技術	42
(1)	文法理論, 機械翻訳	42
(2)	自動分類, 自動索引, 自動抄録	43
(3)	質問応答システム, 事項, 検索	44
(4)	アナログ・データの処理	45
VI	ネットワーク	47
(1)	トリ(樹型)ネットワークとメッシュ(網目)ネットワーク	47
(2)	クラブ型ネットワーク	49
(3)	倉庫型ネットワーク	50
(4)	データ・ソースとファイルのネットワーク	51
(5)	ファイルトサービス端末とユーザの間のネットワーク	51
i	サービス用ファイルを持たないサービス端末	52
ii	サービス用ファイルを持つサービス端末	52
(6)	ファイル間ネットワーク	55
i	定常的データ交換によって結ばれるファイル	55
ii	要求に応じてデータ交換が起る場合	56
iii	ファイル間に階層構造のある場合	57
(7)	ファイル・サイズとファイル・シェア	57
第2部 データバンクの事例研究		
I	各事例に対する考察	65
II	カリフォルニア州のパブリックセクターにおけるデータベース	75
(1)	現状と問題点	77
(2)	形成, 運営, 発展に対する考え方	87
III	組織内データベース	97
(A)	ノースアメリカンロックウェル社	97
(B)	ロッキードミサイル・スペースカンパニー	99
(C)	ニューヨーク連邦準備銀行	101
(D)	ボストン小児科病院	105

(E) ペンシルベニア大学	109
(F) 日本開発銀行	109
(G) 日本長期信用銀行	111
(H) 関西電力	113
(I) 野村総合研究所	115
(J) 鉄鋼連盟	116
IV 企業性データバンク	123
(A) アソシエテッドクレジット・サービス	125
(B) R. L. ポーク社	126
(C) ダン・ブラッドストリート社	128
(D) スタンダード・プア・コーポレーション	130
(E) マグローヒル社	132
(F) 紀伊国屋書店	134
(G) 日本の企業性データバンクに関する討論	135
V 公共的データバンク	143
(A) JETRO	145
(B) JICST	150
第3部 データバンクの分析と評価	159
I コンピュータの発展とデータバンク	161
(1) コンピュータ利用の発展	161
(2) ハードウェアの開発とその普及	165
(3) ソフトウェアの発展	168
(4) コンピュータ技術の発展とデータバンク	170
II 機能的側面からみたデータバンク	174
III データバンクの成立パターン	179
IV パブリックセクターにおけるデータバンク	183
むすび	189
一わが国におけるデータバンクの発展と展望一	



## 第 1 部 データバンクの理論



## I データバンクの概念

人間が社会を形成し、秩序を維持し、文化を発展させてきたある意味では輝かしく、ある意味では苦しい、その長い歴史は、情報を創造し、相互に交換する働きによって支えられたものである。この情報の創造、伝達、吸収が、人間社会で行なわれている仕組みについて、深く、理論的に、そして実証的に解明されていないことにわれわれが気づいたのは、比較的最近のことである。知識産業、データバンク、シンクタンク等、あたかも新しく起って情報現象であるかのように言われるものは、現象そのものが新しいのではなく、その現象を見るわれわれの視点が新しくなったことと、その新しい視点を必要とする新しい社会的要因があったことと、新しい現象であるかのようにその現象形態を変革させる技術的進歩があったことによるものである。

当研究会において、データバンクを、理論的に、そして実証的に、日本の社会の中で成立し得る可能性の分析を含めて研究するに当たり、データバンクといわれる情報現象を、古くからあるものの新しい視点からの探求という姿勢をとった。

データバンクは、単純に考えるとき、データが何らかの形で集められており、それを必要とした人が、必要なものを引き出せる働きを持っている機関であるように思われる。そのとき、データとは何か、それは如何なる性格をもつ機関なのか、それが如何なる場面で働らくのか、等々を分析しようとするとき、あまりにも多くの様相を持つことに気づくのである。

人間のコミュニケーションの歴史の中で発生した情報を、いつでも取り出せるようにしておき必要に応じて流して行くという働きは、文字や記録手段の未発達な段階では「語部」のように口から口への伝承の形ですでに存在していたし、記録手段の普及と共に図書館の形での長い歴史を持っている。日常生活の中でも、個人メモ、帳簿・帳票、資料ファイル、研究ノート等の、多くの手段、多くの形態においてそれを行なっている。そしてフィルムの出現、録音技術の出現、磁気記録の出現によって、新しい様相を生み出している。そして人間の社会活動の中に滲透するにつれて、enumerableできない程の多くの変形を生み出しているのである。

当研究会には、データという名のもとに

「マニュアルまたは機械的に処理・伝送できる形に記号化されたあらゆる情報」を指すことにし、それを「いつでも取り出せるように編成したもの」を、ファイルと呼ぶことによって、データバンクを、

「あるコミュニティにおいて、データの発生と利用があるとき、その中でデータ需要が起ったことによりファイルと、それを処理する仕組みと、ファイル間のネットワークを通してデータを流通させる機能をもつもの」

という形で、情報流通現象の中での機能の概念によって、これらの多くの現象形態を包括的に捉えようとした。

この考え方は、他の多くの考え方、例えばOECD科学政策委員会でトーマス氏の言う

「データバンクは、異なるユーザに有効な情報を与える機能をもった組織、あるいは組織の一部であり、情報の収集・選択、調整、蓄積、検索、縮集、および評価の機能を持つものである。これはMISとは完全に区別するべきであり、またデータに対する数学的処理については、当分の間除外して考える」というものと本質的な差はない。情報流通機能である以上、MISの中で働くファイル(データベース)の機能は含めてもMIS自身とは区別されるべきものである。しかし後述するように、数学的処理は、その回答編集機能としてシステムの中に組み込まれているものは、吾々の場合、一応包含しているとみなしている。またニューヨーク連邦銀行のG.ヘイドン氏の

「データバンクとは、1つの総合データ・ファイルがあり、それに対し多くのユーザが自由にアクセスできるものをいう。データバンクそのものはデータの集積にすぎないが、それに含まれる各データは相互に厳密に整理コントロールされており、データの重複防止、更新・訂正、機密保護などが効率的に行なわれることを前提とする。

数よりは、マルチ・データ・ベース、ネットワークを考慮する点でより広いものである。米国の情報産業社のR.M.マルチノ博士による

「データバンクは、データを自由に預け入れることも、とり出すこともできる、組織化された複数箇のファイルである。

は、われわれの場合、ネットワーク論の所でいう「倉庫型ネットワーク」を持つデータバンクのみを指すものである。また、本研究会に大きい示唆を与えたカリフォルニア州財務長官特別補佐官、M.F.マクドナルド氏の報告の中にある

「機能的データベースとは、常時接近可能な一群のデータファイル(データ・ライブラリ)のことであり、単一の主要な業務と関連をもつ数多くのサブシステムによって統合された形で利用しうるものである。

データバンクとは、数多くの異なる機能と係るデータ・ライブラリの大規模なプールである。これは計画立案もしくは日常業務目的のために統合された形で利用しうるものであり、また業務とは独立して存在しうるものである。」

に対して、われわれは、1つの業務に附属してあるファイルと、多目的の利用に対してそれらとは独立してあるファイルに群とは、後者が前者の発展形態であるけれども、それがデータベースとデータバンクの違いではないと考えている。「ファイル」とは、アイテムの集りから構成さ

れるレコードの集合という、構成要素からみたものを指し、「データ・ベース」とはそのファイルを利用の目的からみたときの呼称であり、データバンクはそのファイルとを処理する仕組み、ネットワークが形成する「流通機能」のレベルでみたときに把握されるものなのである。例えば人事ファイルは、人事管理のシステムにおけるデータベースである。そして、人事ファイルを通して、必要とされた情報を、活動的に流通させる働きに関与する組織構成要素の集まりがデータバンクなのである。

そのような概念規定の下に、データバンクとしてみられる諸事例を、理論的に、そして実証的に、その成立要因、発展過程を含めて研究してきたのである。

しかし、この概念規定によれば、マニュアル操作で処理するファイルも、流通過程に位置づけられるかぎり、データバンクである。その様相のすべてを把握することをこの研究会では意図していない。

#### 「今後の社会におけるデータバンクのあり方」

を模索することに主目的をおく必要がある。そのために過去を振り返ることは、プロセスとしては必要であるが、その発展過程からみた将来を見通すことが必要である。データバンクの発展は、データ発生源の変化、社会におけるデータの需要の、量質における変化によって起るものであるとは言え、現象形態の急速な変化を生んだのは、技術の進歩である。データの需要は、潜在的には常に高度なものを要求しており、それをみたすデータの発生、供給があったとき、需要の高度化が顕在的になるものである。そしてデータの発生は、需要の顕在化が契機となって、新しい、より高度なデータが生まれる。そのシーソー・ゲームの中心にあり、両者のレベル・アップの原動力となっているものがファイル処理やデータ伝送の技術である。

この意味において、今後のデータバンクのあり方を論ずるには、現在の最も高度な技術により処理されるファイル、すなわち、コンピュータライズされ、種々の伝送技術によってリンクされたファイルによるデータバンクを研究の中心対象としたのである。

## Ⅱ データに対する要求の変化とデータ発生源の現状

### (1) データのもつ特性

データは、人間の知的な活動と共に常に造られ、吸収され、その相互の間で交換される情報の記録された形態である。情報量が、それを受けとる前と後で、確率で扱えられる受信者側の構造変化の負の対数で測定されるように、データの効用も、それを利用することによって受信者側に起る変化の大きさが対応する。そのデータ利用による変化が、その利用者に何をもたらすかが、データの価値を律することになる。

得られた情報は人間においては何かの新しい行動を生み出すため（意志決定）に用いられる。直ちに行動を生み出さないときは、記憶の中に止まり、忘れ去られない限りは新しい情報と組み合わせられて行動を生み出す潜在力として貯えられる。そして「物」と異なり、利用頻度の減衰によるライフ・タイムはあるが、廃棄・忘却されなにかぎり消滅することはない。そして情報は人間の生産活動（或いは技術）を通して「物」の上に体化（*embody*）される時、その「物」に働きを与える。例えば部品の集まりに対し、一定の手順で秩序を与える（すなわち設計図により組み立てる）とき、「物」のレベルでしかなかった部品の集まりに、人間が制御・利用できる働きを与える。

また情報は、単独のメッセージで利用されることは、人間の場合には少ない。コンピュータにおいては、単一の“ADD”という命令で演算装置が作動するが、人間においては、他の既知の情報との組み合わせにおいて決定を生み出す。これは決定の内容が、入力したメッセージよりはその入力前を持つ知識により大きく作用されることを意味する。むしろ入力したメッセージは、既知の情報の集まりを動かせる契機を造るのである。すなわち、データのレコード当りの効用は蓄積された活動状態にあるデータの量と共に高まるものである。しかし、これは無限に増加するものでなく、人間の演算能力と演算手段の限界によって制限をうける。すなわち、処理能力以上の情報蓄積は、効用を減じはじめる。このことから、効用とデータ量の相関は、飽和曲線的に扱えることができよう。

このような情報の持つ特性からみて、情報が人間にもたらすもの、すなわち価値を定性的に論ずることができる。

1. 他人の知らないデータをより早く入手することは、他人がそれを知るまでの間に生み出す行動がもたらす相対的利益をもって価値を測定できる。
2. より多くのデータを入手することは、そのデータ蓄積の大きさによる、各レコードの効用の大きさもたらす相対的利益により測定できる。

### 3. より良いデータ

の入手は、それが呼び出した既知の知識の大きさと、その組み合わせによって得られる行動がもたらす相対的利益によって測定できる。

### 4. より正確なデータ

は、それが呼び出す既知の知識におけるノイズを減

らし、誤まった行動を選択する確率を減らすことによる損失の減少の期待値により測定できる。

などである。しかし、一方の特性として情報は、

1. プロセスを経る度に、何らかの歪み（エントロピーの増加）をうける。
2. 情報は「物」と異なり、それ自身が発散的であり、非局所的である。
3. 情報は、それと同内容の既存の知識があるとき効用はゼロであり、また新しく代替情報が発生したとき、効用は激減する。

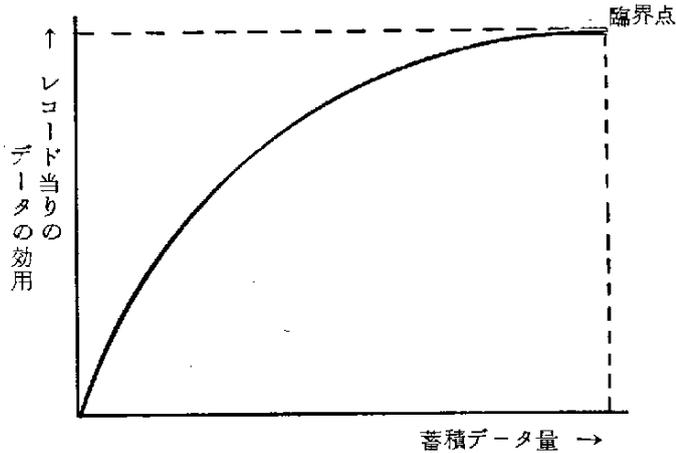
などを持っている。これにより、秘密の保持がむづかしく、ノイズの発生が実に鋭敏（少しの不注意で大誤報をまねくなど）であり、デマゴギーを生じやすい。また、知的生産活動の活発化と共に、ライフ・サイクルが極端に短くなる。

社会は、人間の知的活動によって維持されているものであるから、知的活動の活発化において情報の伝達機能の果たす役割はきわめて大きいと同時に、その情報の流過程が供給主導型であるとき、その供給に偏り、故意にせよ不注意にせよ誤りがあるとき、社会の安全をおびやかすものである。このことは「需要主導型」の流通機構が、効率よく社会の中心に発達しなければならないことを示している。

## (2) データ需要の変化

概念規定でも述べたように、流過程の中で機能的に働くファイルは、長い人間の歴史と共に

第 1.1 図 データの効用と蓄積データ量の相関



常に存在していた。それを、現在の時点において「データバンク」という比較的新しい、まだ成熟していない名で呼び、新しい情報現象であるかのように扱う必要性を考えてみる必要がある。これは、概念規定において強調した技術の進歩が間接的原因ではあっても、そのみでは説明できないものである。

これには、次の問題を指摘することができる。従来、意志決定の段階で、判断材料とするものが少なくとも、それを経験によってカバーして決定を行なっても、危険が少なかったし、データ処理技術も未熟で、その経験による判断を乗り越えることができなかった。しかし、

1. データ処理技術の発展
2. 競争の一般化と共に、誤まった判断のもたらす危険の増加。
3. 商品な技術のライフ・サイクルの短縮と共に、判断業務の頻度の増加とそれに要する時間の短縮。
4. 経営の長期見通しの必要性から起る判断要素の増加。
5. 諸種の業務の内部関連の増加と、境界的、或いは多くの業務の横断面を見通す性質の管理業務の増加。

というような要因が複合的に作用して決定過程が経験には頼れない段階となった。そして、特に社会現象を解明し、現実的な措置を急速にとらなくてはならぬ、多くの問題をかかえている場合には、それが人間を含む現象であるだけに、自然科学的な「実験」は許されない。その実験の失敗によって多くの人間の不幸をまねいてはならないからである。

このような場面では、

1. あらゆる可能性を検討し、
2. あらゆる可能な措置に対し、その結果を定量的に予測し、
3. シミュレーションによって検証する

という過程を経なければ、不用意な決定はしてはならない。

このような立場に立つとき、その決定に必要なデータの網羅的収集が必要である。そのような社会的インパクトが、データをいつでも利用可能にしておくファイルを必要とする方向を生む一つの要因である。

しかし、この角度からのみのデータ需要の切り込み方は、本研究が例証するように、観念論的誤まりを起す危険がある。それは、このような需要は、プランニング、決定レベルへのデータである。この利用は極度に加工度が高い抽象度の高いものである。これに向かったデータの完全な供給は、一挙には到達できない。すなわち、到達すべき地点を示す「旗印」ではあっても、今すぐ踏み込むべき地点ではない。それは、そのようなプランニングや決定を支える、オペレーション

・レベルの仕組みがなければならないからである。そのために、オペレーション・レベルでの仕事から発生し、利用されるデータの利用システムから着実に、ファイル・システムを組み上げないと、ドグマの深い陥穽に陥りかねないのである。これが、M.F. マクドナルドが、その長いカリフォルニア州での経験を通して述べた、彼の言う所の「データベースからデータバンクへ」なのである。

そのような視点に立って、特に社会現象の中で、それを解明したり、対策を決定するのに必要なデータを振かえてみると、大別して次の5種のものがある。

1. 現在供給されており、よく利用されているもの。

これに対してはデータバンクは成立しやすいし、そのファイルに構成・維持のコストと、購買力のバランスの中で、企業的に成立する条件をもっている。この中で、民間で充分、商業ベースで収集し、供給することのできるものでありながら、官庁がそれを行なっているものも数が少なくない。これは情報産業の育成からみて、早く民間に移行すべきである。

2. 現在供給されながら、利用されていないもの。

官公庁の統計の中で、この種のものが多い。これは、供給の方法を変えるか、早く廃止すべきである。少なくとも、この種ものは、自由競争の中にある民間サイドの事業であれば、すでに淘汰されている性質のものである。

3. 発生はしているが、供給されていないもの。

これが企業や個人の秘密に属することであれば、当然のことである。しかし、官公庁内発生データの中に、この種のものが多い。官公庁が、何かの目的で利用したあと廃きしたり、ほごりに埋もれて取われているデータの中に、例えば、何川の水質調査や住民の健康状況のデータなど、今ならば「公害情報」として必要なものが無数にあるのである。

これを流過程に乗せることは、公共機関の、それ自身の業務のためにも、また民間のデータ需要のためにも、大きい責任である。

4. 必要であるとわかっていても、発生すらしていない。

これも官庁の統計などの中に多くの事例がみられる。統計の記入フォーマットの悪さや、集計法のまづさによって、少しの工夫で集まるべきデータが集まらないのはこのケースである。

またこのケースの1つとして、最近起ってきたシンク・タンクも、これを解決するためのインパクトから生まれたものといえよう。

これらを通して言えることは、データの発生の仕組み、収集の仕組みに対する、基本的反省が必要であり、発生そのものが、需要のレギュレーションを受けるべきことを示している。

### (3) 公共財データの発生と流通の現状と問題点

以上のような需要の変化と、需要例からみたデータの発生・供給現状から、先ず公共財データの発生・流通を、事例的に概観してみよう。

#### (i) 発生と流通の現状例

公共財的性格の情報として、磁気テープ化されているものに関する調査結果の代表例として、主要11省庁における約120種類のファイル(工業技術院、電算機利用に関する技術研究会—ファイル研究班、報告書参照のこと)をみると、大部分が、各省庁個有業務処理に当たった電算機システムのデータ・ベースそのものである。

これらのファイルは、極く例外を除いては、フィジカルな条件のコンバートは可能になっており、ロジカル・レコードの互換性については、目下、上記「利用研」で検討が進められつつある。

利用に当たっては、このデータ・ベースと、処理プログラムのコンビネーションによってこれらの情報の活用は、可能であるので、データ・ベース間の相互交換ネット・ワーク形成によって、種々のデータの組み合わせも出来るので、このような、個有業務のデータ・ベースと、各データ・ベース間のネットワークの形成により、政府機関におけるデータバンク機能は、成立できるわけであり、電算機および通信関係の現在の技術の延長線上に、このようなデータバンク機能成立のためのボトルネックになる要因は見当たらない。

また民間との関係についても、各個有業務は、民間情報のインプットから出発しているので、アウトプットの面にオープン・エンデッド考え方をとれば、民間利用も可能になってくる。

例として、政府機関における電算機適用業務に関する調査(行政管理庁、管理局調)をみると全体の業務(170台、レンタル費83.6億円)に対し、集計計算型の適用が39.3%管理型の適用が35.6%、分析予測型の適用が10.6%、試験研究型の適用が14.5%になっており、この集計計算型に使用される年間延べのマシントイムは約25,600時間の多きにのぼっている。このうち、データ交換が行われているものは(情報処理開発センター調査46.2月)、54例もあり(大部分が磁気テープによるオフライン方式である。)、そのサイクルは月単位と年単位が大部分であるが、1サイクル当りのデータの延件数は、週単位以下約2,718,000、月単位約1,276,000、年単位約70,000となっている。これらの交換データは現在のところ、オペレーショナルなファンクションに関するものが大部分であるが、一部には、プランニングファンクションに関するものも含まれており、ネットワーク形成は、必要に応じて徐々に進行しているのが現実である。

#### (ii) 公共財情報の問題点

以上あげたように、公共財情報の流通をさまたげる要因は、技術的には小さいといえるが、

次にかかせるような、データそのものを中心とする問題点があり、今後の課題になっている。

### 1. 各省庁個有業務結果の活用

個有業務の電算化は、夫々単独完結的に最適条件を追究して行われて来た。利用技術の標準化を更に押しすすめて、処理結果を流通し易くすることが、全体に対する最適化へのアプローチの第一歩である。

### 2. 統計調査への柔軟性の要請

統計は、本来連続性を重んずる上から保守性をもつことは、或る程度まぬがれ難いが、同時に急速に変化する社会・経済現象に対し、カバレイチにおいても、内容においても対応性がなければならぬであろう。 $\frac{1}{4}$ 世紀前には考えられなかった領域の拡大（例えば医薬保険関係）、当時は非常に詳細、正確性を要求された内容の変化（例えば物資関係）等につき、対応性が小さい。特に、最近では電算機の各種業務適用の拡大に伴って発生するインディケータとの代替（労働移動関係）のような新たな現象も出て来ている。収集、利用の両面にわたり、技術的にも制度的にも柔軟性の不足が感じられる。

統計の多角利用については、調査の限点から、データ、リンケージを可能にすることが必要であり、その基盤として、他のレジストレーションから発生するインディケータと接続可能にする統計側から歩みよりも必要であり、それは必ずしも統計の中立性や調査対象の秘密保護を損う問題にはならない。

### 3. 新領域のインディケータの開発

需要の側面において、今後のビジョンとして、共通に指摘されたのは、産業の情報化への指向である。企業間の競争次元が、従来は技術力、資本力の面に加えて、新しいシステム化のためのソフトな技術力を含む総合経営力の面を加えることになり、また、公共部門においても、社会環境問題のクローズアップによる単論理的思考から多論理思考へのシフトが大巾に行われることになろう。

情報需要は益々多様化し、未開拓の領域のデータ開発を強く要求してくることになり、従来は価値尺度では、数量化困難な現象に対する新メジャーによる数量化技術、アンコンテナヤスなデータの評価技術等を開発する必要がある。

### 4. プライバシーの確保

供給側における機能が高まるに従って、各省庁個有業務や、統計処理業務等のデータベースで留意されなければならないことは、個人の秘密に属する情報の扱いである。これは電算機導入による以前と本質的には、変るところがなく、当事者の職業倫理の問題である。電算機利用によるファイル管理のマニアルの完備により、むしろ事故の防止は、以前よりも容易になるという考え

表 1.1 情報処理

手 法	業 務	イ 生 産	ロ 資 材 購 買	ハ 在 庫 (製 品)	ニ 原 価	ホ 販 売	ヘ 輸 送
1. 統計表作成(グラフ相関表など)		{ 54 3	36 1	47 5	31 4	100 19	11 0
2. 統計的分析(回帰分析など)		{ 44 4	14 1	20 1	10 0	47 8	8 1
3. 数学的計画法・LP		{ 46 6	11 1	10 2	5 0	12 3	19 2
4. その他の数学的計画法(QP・DP)		{ 14 0	2 0	3 0	3 0	1 0	1 0
5. ネットワーク(PERT・CPM)		{ 30 3	2 0	0 0	0 0	2 0	1 0
6. スケジューリング		{ 31 8	7 0	5 1	0 0	0 0	0 0
7. エコノメトリックス		{ 4 0	0 0	1 0	2 0	4 1	0 0
8. サイコメトリックス		{ 0 0	1 0	1 0	0 0	0 0	0 0
9. ソジオメトリ		{ 0 0	0 0	0 0	0 0	0 0	0 0
10. テクノメトリックス		{ 1 0	0 0	0 0	0 0	0 0	0 0
11. エンジニアリングエコノミ		{ 7 2	2 0	2 0	2 0	1 1	1 0
12. シミュレーション		{ 47 4	13 1	29 3	5 1	15 2	19 1
13. その他のOR手法		{ 11 0	9 0	13 3	1 0	9 1	2 0
14. リアルタイムデータ処理		{ 5 0	1 0	5 0	0 0	5 3	0 0
15. インフォメーショントリバル		{ 7 1	1 0	2 0	1 0	2 0	0 0
合 計	総 数 内◎印の数	{ 301 31	99 4	138 45	60 6	198 38	62 4

注： ◎印は、特に重点をおく場合

利用手法（現在）

ト設備	チ経 理	リ給 与	ヌ人 事	ル財 務	ヲ総 務	ワ調 査	カ情 管 理	ヨ全 管 理	合 計
11 0	35 3	69 4	50 1	21 1	7 0	40 0	13 0	12 1	537 42
6 0	10 1	9 0	14 2	7 1	0 0	41 2	6 0	3 0	239 21
14 2	4 0	1 0	2 0	3 0	0 0	11 1	2 1	1 0	141 18
7 1	0 0	0 0	1 1	1 0	0 0	3 0	0 0	0 0	36 3
37 3	0 0	0 0	0 0	0 0	0 0	6 0	0 0	0 0	78 6
3 0	0 0	0 0	0 0	0 0	0 0	4 0	0 0	1 0	51 9
1 0	4 0	0 0	2 0	1 0	1 0	20 5	0 0	2 0	42 6
0 0	0 0	0 0	1 0	0 0	0 0	0 0	0 0	0 0	3 0
0 0	0 0	0 0	0 0	0 0	0 0	0 0	0 0	0 0	0 0
1 0	0 0	0 0	0 0	0 0	0 0	0 0	0 0	0 0	2 0
7 0	1 0	0 0	0 0	0 0	0 0	1 0	0 0	0 0	24 3
28 1	10 0	5 0	11 0	7 0	0 0	11 2	3 1	8 0	213 16
4 0	0 0	1 0	3 0	1 0	0 0	4 0	0 0	2 0	60 4
0 0	0 0	0 0	1 0	0 0	0 0	0 0	0 0	0 0	17 3
2 0	0 0	1 0	5 1	0 0	1 0	5 0	11 2	2 0	40 4
121 7	64 4	86 4	89 5	42 2	9 0	146 10	35 4	21 1	1,480 135

表 1.2 情報処理利用

手 法	業 務	イ 生 産	ロ 資 材 買	ハ 在 庫 (製 品)	ニ 原 価	ホ 販 売	ヘ 輸 送
1. 統計表作成(グラフ相関表など)		{ 33 1	31 0	38 2	24 2	54 5	7 0
2. 統計的 分析(回帰分析など)		{ 36 1	17 1	25 1	22 1	73 1	8 1
3. 数学的 計画法・L P		{ 87 15	29 0	44 3	15 1	39 5	50 5
4. その他の数学的計画法(QP・DP)		{ 29 1	8 0	20 2	6 1	13 1	9 1
5. ネットワーク・(PERT・CPM)		{ 58 5	7 0	5 0	2 0	10 2	8 0
6. スケジューリング		{ 60 18	17 4	10 2	0 0	8 1	12 0
7. エコノメトリックス		{ 5 0	3 0	2 0	7 0	19 1	2 0
8. サイコメトリックス		{ 0 0	0 0	0 0	0 0	7 1	0 0
9. ソシオメトリ		{ 0 0	0 0	0 0	0 0	4 1	0 0
10. テクノメトリックス		{ 5 0	0 0	0 0	1 0	0 0	1 0
11. エンジニアリングエコノミ		{ 21 1	4 0	3 0	9 1	3 0	2 0
12. シミュレーション		{ 76 13	27 0	54 9	16 1	48 9	45 5
13. その他の O P 手法		{ 16 0	13 1	19 1	6 0	15 2	10 0
14. リアルタイムデータ処理		{ 39 12	20 1	54 11	4 2	48 13	8 1
15. インフォメーションリトリバー		{ 7 1	4 1	5 0	2 0	12 2	0 0
合 計	総 内◎印の 数	{ 472 81	180 10	277 35	117 9	352 61	153 15

注： ◎印は、特に重点をおく場合

手法（今後の予定）

ト設備	チ経理	リ給与	ヌ人事	ル財務	ヲ総務	ワ調査	カ情報	ヨ全般	合計
70	310	380	302	170	70	301	121	152	37416
81	100	161	171	91	10	382	131	132	30622
171	90	20	111	90	00	151	80	121	34733
121	30	10	30	50	00	50	10	80	1247
505	00	00	20	20	00	30	50	101	16213
71	00	00	20	20	00	30	31	40	12827
30	41	10	31	62	20	284	50	60	969
00	00	11	80	00	00	11	10	00	183
00	00	20	00	00	00	11	00	00	72
20	00	00	00	00	00	11	10	00	111
223	00	00	10	30	00	21	10	30	746
399	203	50	272	255	20	224	80	379	44169
110	40	10	30	30	11	101	51	132	1299
10	40	40	30	30	00	20	214	122	22446
10	31	10	291	32	20	200	6111	164	16623
18022	877	702	14010	8612	161	18321	14425	15029	2607340

方で対処すべきであるが、プライバシー問題から、ともすると情報提供側が定めた処理（例えば集計）のみによるサービスに限る傾向が、官庁統計になしとしないが、過剰措置は今後のビジョンにはなじまない。

#### (iii) 処理手法とデータ供給

情報処理手法の民間における適用状況の調査（68年調査、約3,000対象アンケートによる）により、表1（現状）、表2（今後の予定）を得たが、統計表の作成に、統計的分析などの基礎統計の処理は、こなし終えて、LP、DP、QP、PERT、GPM、スケジューリング、シュミレーションなどの数学的計画法などを用いた方法に重点が移行しようとしており、またリアルタイム処理、インフォメーション・リトリバー（データ・バンク機能に対応）などの情報処理機能の高度な手法の導入が、積極的に予定されていることがわかる。

これに反し、エコノメトリックス、サイコメトリックス、ソシオメトリックスなどの計量科学的手法の面では、今後も大巾な採用予定がなく、その理由にモデルビルディングの困難さとともに、ソーシャルインディケータの不足をあげたケースが多いことを見逃せない。

適用業務は、生産、販売、輸送、経理等の定型業務から、調査、情報管理、管理全般が加えられる予定が目立ち、オペレーショナルファンクションの適用が進むにつれて、プランニングファンクションへの拡張が予定される傾向は明白によみとれる。その後の追跡調査が望まれるが、計量科学的手法の普及は、インディケータ不足のために停滞しているという結果を予測することは難かしくないであろう。

#### (iv) 情報処理システムに要求される特性

##### (1) オープン・エンディッドな性質

オーガニゼーションの活動は、具体的には形態が異ってみえるが、基礎的な機能としては、1つは、意志決定に連なるプランニングの機能、一つは、この決定に従った行動（リアライズ）と統制（コントロール）から成り立っているとみることができる。

これらの機能は、情報の伝達と処理という手段を媒介としてその機能を働かせている。

このような情報機能は、次のようにブレイクダウンさせることができよう。

- (a) 資料の認識（内部、外部からの資料のソースや入手ルートの問題 含フィードバック）
- (b) 資料の評価（受信の意味作用）
- (c) 情報の記憶（体系化の蓄積 含廃棄 検策）
- (d) 情報の加工・分析・結果
- (e) 情報の伝達（含社外との交換）

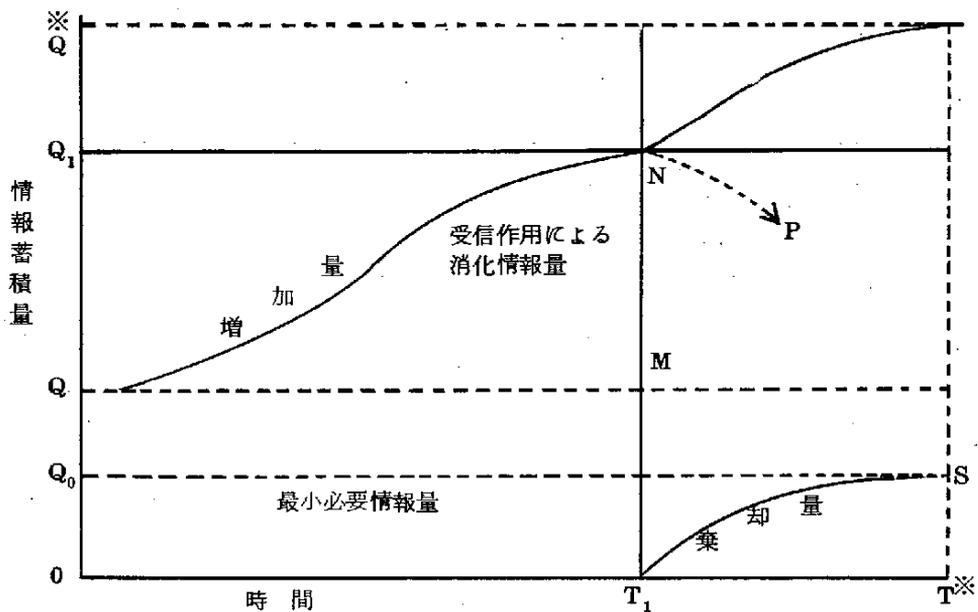
の条件をみたせば一義的に定りやすく、「閉じた」システムとして考えることもできるが、(a)~(d)は、一義的に定めることは極めて困難であり、マン・マシンインタラクションによってカバーされなければならない「開いた」性質のものもある。

(2) スケールメリットが保証されない性質

大量、単純な業務の機械化というケースに限った電算機による情報処理は、定型化可能な低次の頭脳労働（一種の肉体労働ともみることができる。）の場合は、スケールメリットは期待できるが、総合経営力の競争次元への参加や単理論から多理論への思考のシフトのような場合には、プランニングファンクションやコマンドファンクションにおける情報処理システムにスケールメリットは保証されないと考えることができよう。

受信の意味作用は、第1図のように、情報の蓄積ベースがあってはじめて成り立つものであるが、受信の意味作用の深さ（情報の評価力）は、蓄積情報量（過去の学識経験）と何等かの相関がある（一説には積分関係で説明可能ともいわれている）と考えられ、システムとしての合目的な情報が選択され得る。しかし、やがて情報の増加が、システムの容量に対して飽和状態に近くなると、このような受信作用には不消化現象を生ずることになり、蓄積情報の棄却能力を働かせることを必要とすることになってくる。

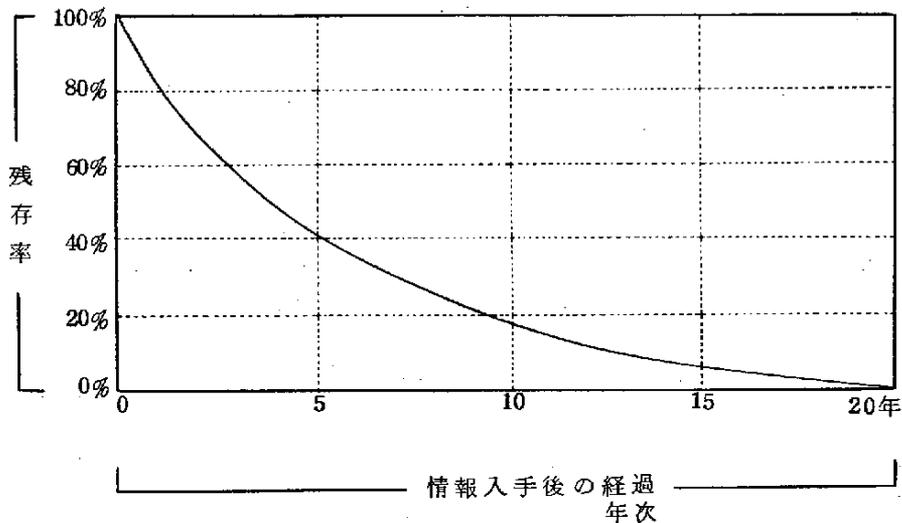
第 1.2 図 受信作用と情報蓄積量との関係



(注) 情報処理機能の自己発展には最小必要情報量(O T, M Q)が不可欠である。その上に受信作用による消化情報が蓄積される(Q M N)。しかし $T_1$ になるとシステムとして飽和点Nに達し、自己発展はとまる。そこで $T_1$ から $T_2$ 、 $T^*S$ 部分だけの情報を棄却していくと情報の増加量曲線は上方にシフトし、 $T^*$ では、飽和点は $Q_1$ から $Q^*$ に移行する。もし $T_1$ で不要情報を棄却しないと増加量曲線は不消化現象によりNP方向に漸減し、システムとしての機能は減衰する。

飽和点と、棄却についての一般論に関する研究は、なお続けられているが、経験的には、図2のような、時間経過と、蓄積情報の価値の減衰の相関が確認されている。(図2は防衛庁の戦略情報の場合で、有効蓄積期間は6年になっている。)

第 1.3 図 有価値情報の残存率の推移



このような結果に、原子力研究所、通産省データセンターでも確認されており、蓄積、棄却、受信の意味作用の関係から、システムの発展方向は、分化、専門化の度合いを強める程、高密度化し機能がよくなるといえよう。

従って、オール・オーバーな領域を対象にした集中的なシステムは、極めて非現実的であると考えられる。

(3) ネット・ワークの形成の意義

個有業務毎のデータ・ベースの成立や、プランニングファンクションに連がる高密度なデータ

・バンク機能の分化、専業化への展開が現実のものとなれば、これらの個有システムのファイルにアクセス可能で、自由にデータを組み合わせて活用出来るような、高度の選択機能を保証できる情報の流通機能が真のデータ・バンクに外ならない。このような情報のネット・ワークの持つ意義を整理しておく必要がある。(純技術的な問題は、第4章で述べることにする。)

データ・バンクの理想は誰でもが、何時でも、どこからでも、欲しい情報を入手出来るという条件を満たすことであって、それはシステムの内外をとわず誰でもが参加することができ、情報の選択・流通に不用の制限を加えないということである。

情報の価値は、一般的には、正確性、速報性、網羅性等の条件によって定まるが、加工・分析における知的創造力によって附加される価値が最も大きいことから、このような知的創造力を損なう条件がないことが望ましい。

このような情報の選択・流通の条件を満たすための情報ファイル間の接続体系には、ネットワーク体系が適切であろう。

現在、われわれは、マスコミュニケーションのような一方通行の情報流通の体系や、組織機構上の上・下の関係に重きをおいたコミュニケーション体系に多くなじんでいるが、このようなツリー体系(ピラミッド構造のような)に、情報ファイルが入った場合は、限られた条件(例えば個有業務)に対しては、能率がよいが、参加・選択の自由度は著しく制約されることになり、多彩な情報の組み合わせによる知的創造性の発揮は難しくなってくる。

以上のような意味で、情報のネットワークの持つ意義は、今後も各方面で充分吟味されることが望ましい。

それは、国全体の中核管理機能に影響するような組織原理につながる問題であり、思想的には、合理主義よりも経験主義をとり、技術的には、能率至上よりも安全を指向し、経済的には、競争原理を大巾にとり入れ、組織的には、ハード構造からソフト構造への転換を意味するからである。

### Ⅲ データの収集の分析

データバンクの現状、将来を考えるに当って、データの発生と需要の高度化のシーソー・ゲームの支点となるものが、技術の発展であるとした。その論点に立脚して、この節につづくいくつかの節において、データの収集・分析・処理の技術を概観し、問題指摘を行なう。

データバンクの動きを、コミュニケーションの技術としてみると、

「時間的にも空間的にも、またその意図する所においてもランダムな多量のデータ発生があり、一方個々のユーザーの利用目的や要求内容の異なる需要が時間的にもランダムに多数存在するとき、その間を結ぶコミュニケーション・チャネル」

と考えることができる。

このコミュニケーション・チャネルの特性は、次のように捉えることができる。

(1) 従来のコミュニケーション・モデルのように発信から受信に一方的に情報が流れるのではなく、

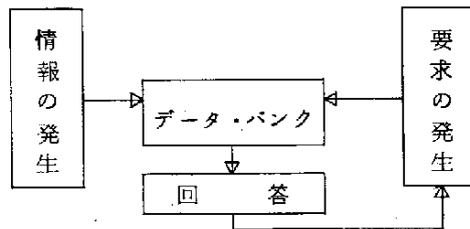
○情報という工程を窓口として入って来る情報と

○質問受付という工程を窓口として入って来る情報とを

○比較・照合して情報を選別し、質問情報の示す仕様によって処理して、回答編集という窓口を通して出力する。

という、2つの入口と1つの出口をもつ。(第1.4図)

第1.4図 データバンクのコミュニケーション・チャネルとしての特性



(2) 情報の発生時点と要求の発生時点に時間差(例えば特許資料は、有効期間が17年であるが、期間をすぎたものに対する要求も相当量ある。)があるため、ファイルを必要とする。

この事から、このコミュニケーション・チャネルの動きは、次のサブ・ファンクションを持つ。

1. ファイル構成と維持
2. 質問の分析と構成
3. 質問をガイドとするファイル探索
4. 回答の処理と編集
5. 情報の交換と伝送

先ず最初に、この中の収集と分析の技術から概観を開始しよう。なお、この章を通して多くの技術用語が用いられる。その用語の解説を、現われる度に行なうのは、記述を繁雑にさせるのでこの章の終りに定義をまとめてあるので、疑問とと思う用語の現われたときはその定義を参照されたい。

## (1) 収 集

ファイルを構成する最初の工程は収集である。何らかの利用目的にとって、ファイルがデータベースとして有用であるか否かの最も明確な判定基準は、必要とするデータが収集されているか否かである。この事は言い換えれば、特定のデータを収集するか否かの判定基準は、利用目的の方から定められるものであり、無目的に収集したデータ・ファイルはそのままではデータベースとはなり得ないし、また、あらゆる目的に適合する1つのデータベースも、架空のことでしかない。

収集という工程は、ファイル構成側が、データ源に対して加えるオペレーションである。データ源は、能動的データ源とみなし得るものと受動的データ源とみなし得るものとに大別できる。

能動的データ源： 研究論文や調査報告、定期的に提出される報告書、官公庁の登録業務を通して提出されるデータなど、自発的または制度を背景としてデータを供給するデータ源。これに対しては、受動的収集オペレーションが対応する。

受動的データ源： アンケート、面接、調査員の巡回、調査活動など、能動的収集オペレーションを加えることによってデータが供給されるようなデータ源。米国の信用情報のデータバンクでは、数千人の調査員がデータ収集業務にたずさわっている。

収集工程は一般に多額のコストと人的資源を必要とする。特に受動的なデータ源に対する収集工程においては、ファイル構成コストと労力の大半が、収集工程に費やされると考えてよい。

データ・バンクが情報産業の一分野として成立するためには、収集コスト、労力の低減が最も大きい課題である。そして現実にはデータ・バンクとして成立し得ているものは、

- (1) 比較的収集コスト、労力が少なくすむデータの場合

- (2) 収集コストを越して大きい売上げの期待できる性質のデータの場合
- (3) 他の仕事の副産物としてデータが収集できるため、収集コスト、労力のすべてがデータ・バンク側の負担とならない場合
- (4) 公的な資金によって収集活動が行なわれ、豊富なデータが安価に供給される場合のどれかに属している。

この収集コストと労力を低減させるための技術的な問題点を列記すると次のようになる。

(i) ファイル構成者において留意すべき点

1. ファイルの利用目的（現在ばかりでなく予測できる将来の目的をも含めて）をよく分析してレコードを構成するアイテムを選択し、バリューの抽出基準を立てること。

2. 各データ源について、それから得たレコードの利用頻度、有用度の統計をとり、データ源の評価を行なうこと。

3. データ発生即データ収集（入力）となるようシステム設計をすること。帳票設計におけるワン・ライティングは普及しているが、原始帳票への記入即マシンリーダブルな入力という方式はまだ普及していない。これはMICR、OCRなどのハードウェアが普及していないことにも原因があるが、それよりも、

(a) 発生したデータは、必ずファイルを通してのコミュニケーションに乗せられるべき価値を持っていることに対する認識の不足。

(b) データを発生させた部局では必ずしもマシンリーダブルでなくてもよい場合も、必ず他の工程において利用価値を持っており、そのためにはマシンリーダブルにしておくことが必要であるという認識の不足。

(c) 帳票の設計者に、マークセンシングやOCRの知識が少なく、転記（コピー、パンチも含む）を中心としたコミュニケーションを前提として設計すること。

などの方が、大きいボトル・ネックである。

(ii) データ源において留意すべき点

1. 能動的データ源においては、データを供給する際に、そのデータのより効率よい流通に留意すべきである。例えば学術論文の著者は、その論文内容をよく表現する標題と、著者抄録と検索タグとなるキーワードを必ずつけることを習慣づけることなどは、この一例である。

2. 受動的データ源においては、情報社会の一員として、情報が「財」であることをよく認識すると共に、その財を利用し得る権利を持つと共に、それを供給する社会的義務をあわせ持っていることを自覚すべきである。

### ④ 公共機関の責任

1. 官公庁および公共的性格をもつ機関は、その業務を通して、大きなデータ発生源である。しかし、わが国においてはそのデータを、行政等その直接業務の目的にのみ使用し、公共的な「財」であることを殆ど意識していないことは、極めて残念である。情報社会における公共機関として、その中で発生するデータを社会に供給する責任を自覚すべきである。

2. 官公庁は、民間の情報産業の育成のためには、イニシャル・コストやランニング・コストの負担、金融措置ばかりでなく、それよりも、民間機関ではコストにおいても労力においても負担することが困難なデータ収集活動および公共財としてのデータの永久保存を行なうべきである。必要とあれば、データ収集・供給・保管のための公共企業体を設立すべきである。

3. 公共機関が、その持つデータ源をもって、みずからデータ・バンク機能を果たすことは、民間の情報産業を圧迫する危険がある。むしろ、各省庁や地方自治体で発生するデータに対し、

(a) 機密やプライバシーに属するものは除き、すべて国民の利用し得る財として、印刷物およびマシンリーダブル形式で供給することを義務づけること。

(b) マシンリーダブル形式で供給するとき、標準フォーマットを設定すること。現在、米国化学会の下部機関である Chemical Abstracts Service では、化学関係の全世界で発表される論文を紹介する活動を行っており、その一部をマシンリーダブル形式で供給しているが、それをすべて ASCII コード (ISO の勧告をベースとした、8ビット構成の、米国標準文字セット) を用いた Standard Distribution Format によって供給している。そして ISO に対し、学術文献のファイルに対する国際的標準フォーマットにするよう提案している。JIS においても、文字セットばかりでなく、諸種のデータに対して情報交換用の標準フォーマットを設定すべきである。

4. 公共機関はデータ供給に当って、機密やプライバシーを損なう場合を除いて、加工や変形をしてはならない。その理由は、

(a) 公共機関で発生したり収集したデータは利用範囲が広いので、特定目的のために加工するとそれ以外の目的のための利用者にとって価値が減少する。できる限り生データである方が望ましい。

(b) 下手な変形を行なうと、データに歪みが生じ、誤った判断を起す危険がある。

5. 公共機関は、どのデータに対し、少なくとも国内のどこにそのデータ源があるかを調査し、

(a) Who, Where, What 情報を供給すること。すなわち、どこが誰が、如何なるデータを発生させているかを教える機能をもつべきである。

(b) データ・ファイルの所在情報を供給すること。すなわち、どこにそのデータのファイルがあるかを教える機能を持つべきである。

(c) スウィッチング機能を持つべきこと。データを必要としたとき、公共窓口に問合わせると、そのデータを供給するバンクにスイッチし、そのバンクからデータが供給される仕組みを構成すべきである。

## (2) 分 析

データの分析は、収集について、各データをファイルに収納するに当りファイルに アイテムの値を決定する過程である。その各アイテム・ヴァリュースが後の利用の対象となるのであるからファイルの品質を決定する大きな要因である。

この分析法は、多くのデータについて、技術的に確立されている段階ではなく、各種データについて、伝統と経験の中悉処理されている。学術的数値データについては、それぞれの学術分野の理論的背景によって分析法は一応確立されているとみられるが、それ以外については、標準化されたパターンはない。このため、分析過程は人間によらざるを得ず、個人差、人為的な操作、錯誤が起り、品質管理が困難となる。これはセンセンシアル・データにおいて、特に顕著である。

この現状において考慮すべき点は次のようである。

1. 分析にあたって、オリジナル・データに対して忠実であること。
2. ユーザ側の要請・フィードバックによって、分析基準が定められるべきであって、ファイル作成者側の主観・独断は避けるべきである。
3. 各種データの分析能力を持つ人材を、社会の責任において養成すべきである。
4. 各種データについて、ファイル構成のアイテム、およびヴァリュースの表記法について、そのデータの発生・利用を含めた関係者の間で標準化を進めるべきである。
5. 情報関係者の間で分析基準設定についての積極的協議が起るべきである。
6. 各種データの分析法、その自動化について、基礎・応用の研究開発を、官民の協力によって公共的に促進すべきである。

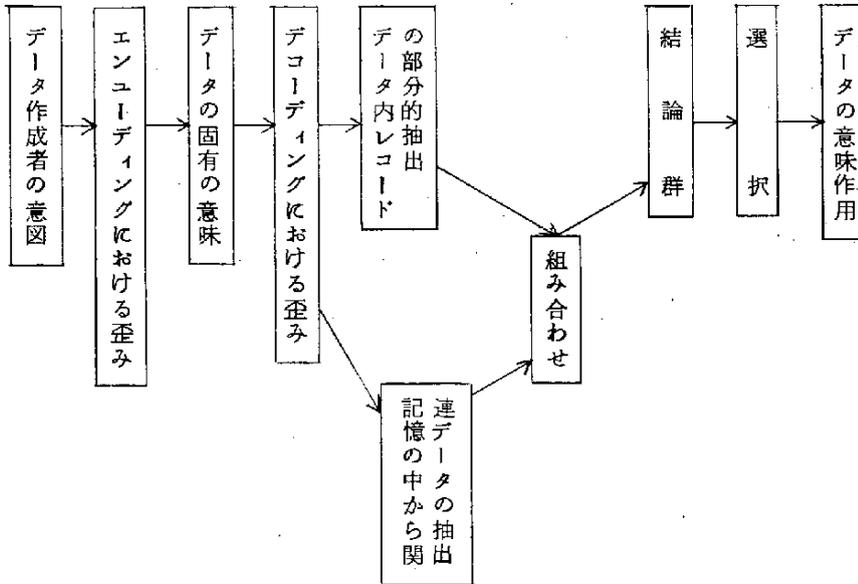
7. 実際の分析において、レコードの有効利用をはかるため（分析コストのパフォーマンスを高めるため）、多くの利用目的に向けて、分析基準を立てる必要が起る。(Single-analysis multiple-use)

この1および2については、次の問題を併わせて考慮する必要がある。それは

「データの利用者は、必ずしも、そのデータの作成者の作成意図、そのデータがそれ自身でもっている意味の通りには利用しない。その利用の方法、利用の内容は、利用者によ

て千差万別であり、それを供給する側では予測できない

ということである。すなわち、データ作成者の意図は「エンコーディングにおける歪み」によって、そのデータ自身が持つ意味（固有の意味）とは異なって来る。



また、それを利用する側において

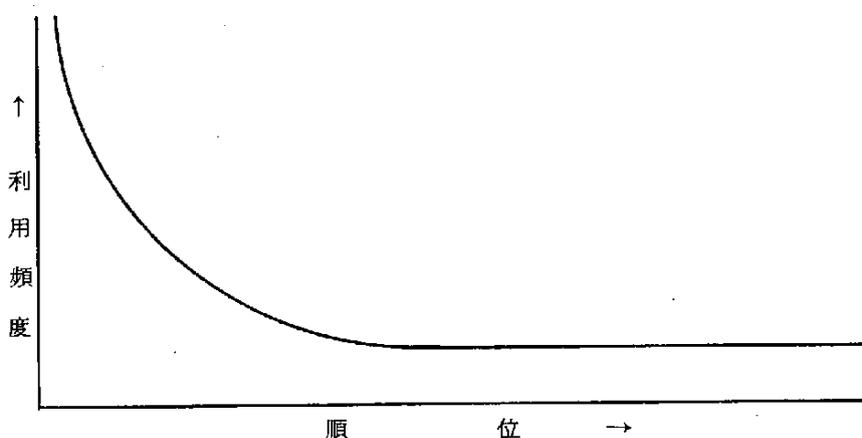
1. デコーディングの歪み
2. データ内容の1部分のみを利用する場合のあること
3. そのデータと関連する既知のデータと組み合わせ、推論すること
4. その得られた結論の中から、適したものを選択すること

という過程を経て、そのデータの受信者に与える意味（意味作用）が得られる。これはこの1～4の過程の差異によって、多くのバリエーションがあり、予測することはできない。このために、分析は、オリジナル・データに忠実であり、余分な操作を加えることは利用者にとってノイズの原因を造ることになるからである。

次に、ファイルの網ら性メリットと、スケール・デメリットの問題がある。利用者にとって、頻繁に用いるデータ・ソース（これはファイル内レコードと言いかえてもよい）とあまり用いないものがある。これを、利用する頻度の多いものからつけた順位と、頻度の相関をとると、次のグラフのようになる。

これは、順位が下がるにつれ、最初激しく減退し、その後同程度の頻度のものが数多く続く形をとる。（大体  $\log - \log$  で直線的）そして、利用者によって、データ・ソースの利用頻度・

第 1.5 図 データ・ソースの利用頻度とその順位の相関グラフ



順位は異なる。例えば化学関係の研究者と電子工業者では、ソースの利用が全く異なる。そして利用頻度の高いその専門分野自身のソース（例えば学術雑誌）の数の10倍以上の他の分野のソースが、個々のソース当りの利用頻度は少ないけれども、データを供給している。その傾向は、今後増大するであろう。

このとき、次の2つの問題が起る。

1. サービス対象を特定ユーザ・グループに限定すると、それにデータを、利用頻度は少なくとも供給するソースを収集し、分析する必要がある。データ量を拡大するとき、利用頻度の高いものから（これは経験的にわかるため）収集・分析する。そして漸次マージナルな効用のものに拡大するため、収集・分析コストが、利用当りで上昇する。このため、パフォーマンス当りのコストは、スケールを増すにつれて上昇することになる。これを「スケール・デメリット」と表現した。

2. しかし、サービス対象を拡大すると、特定のユーザ・グループにおいてマージナルであるデータ・ソースが、他のグループにおいて利用頻度が高くなることもある。このことを「網ら性メリット」と表現した。

これは、データ量を増すときに、マスト・パフォーマンスを落さないためには、サービス対象をも拡大する必要のあることを示しており、運営政策、コスト分析において重要であり、ファイル・シェアを論ずる上でもよく議論する必要がある。

## Ⅳ ファイルの構成と探索

ファイルが情報の流通過程の中で、データ・ベースとして活動的に機能するためには、次のような性質を持たねばならない。

1. そのファイルが対象とする個体に関し、新しいデータが間断なく発生するとき、その属性が時間と共に変化しても、その値をアイテム・ヴァリューとしてファイル内に表現できる構造をもっていること。

2. 質問に対してファイルが答えるためには、サーチャブル・データとクェリー・データとが照合できなければならない。その照合を容易にするタグ記号の管理システムをもつこと。

3. 新しい1次データの入力によって、ファイルが容易に更新できること。

4. ファイルを管理する責任と権利が明らかになっており、その権利または責任を持つ者以外はファイル更新ができない管理システムをもつこと。

5. 他ファイルとのデータ交換が可能であること。<sup>(注2)</sup>このためには、レコードのアイテム構成と、各アイテムにアイテム・ヴァリューを与える基準についての標準化が必要である。また、活動中のファイルの間でデータ交換を行なう場合のインターフェースにおいて、標準的の文字セットと記録コード、標準的ファイル・フォーマットが設定されていれば多くの、これから起るであろうファイル間データ交換のインターフェース設計、更にそのネットワーク設計において、技術的問題の大きい部分が解決できる。

(注2) これは、単一のジョブに対してのみ機能する孤立したシングル・データ・ベースの場合には必要ではない。しかし、情報の流通過程において活動的に機能するためには多くのジョブに対し、多くのファイルがデータ・ベースとして働くことが必要である。或る時点において、単一のジョブにのみ働くファイルであっても、マルチ・データ・ベースの中で機能する可能性を秘めているものと考えることが、常に必要である。

6. データ管理のシステムを考えると、データ利用の仕組みの変化(要求そのものの変化、回答の編集や演算の仕組みの成長など)が流動的であることを考え、現在必要な機能のみに対して徹底的効率を高めることよりも、システム全体の生長をも考慮した巨視的な効率を軸として考える必要がある。特に、シングル・データ・ベースからマルチ・データ・ベースへ、また出力結果に更により複雑な加工ソフトウェアが附加される可能性などは、常に起って来る問題であるからである。

## (1) ファイル構成の技術

単一のファイルを構成する技術は、アルファニューメリック・データに関するかぎりは、現在のハードウェア（すなわち第3世代といわれるもの）に関しては、大体整理が終っていると考えてよい。センテンス・データは、そのデータそのものの研究がまだ未熟であり、アナログ・データに関するその取扱いについての研究が進行中であるため、本質的には多くの問題があるが、それをアルファニューメリック化したものについては、同様である。

### (i) データ交換用ファイル

まず、1つの活動中のファイルから他のファイルへデータを移動するための、中間媒体となる磁気テープ・ファイルのフォーマットは、大体の所、ダイレクトリー方式の可変長形式をもつシリアル・ファイルで、記録コードはISO勧告をベースとした標準コードによるものに向う傾向が強い。これは、MTによる情報の交換、販売の問題であり、従来の出版形式に加わる、強力なコミュニケーション手段として、重要な問題である。最近の顕著な例としては、学術的2次情報サービスにおいて、印刷物と併行したり、印刷物に代わってMTによる供給（リース形式が多い）されるケースが多い。この中には、データの供給のみもあるが、ソフトウェア・パッケージも併わせて供給する所も多い。

この分野でも米国の主導性は歴然としており、米国内のデータ交換のための標準化が進行中である。例えばコンピュータの利用者の中での標準化運動がある。また米国の発言により国際的標準化への働きかけも活潑である。これに対し、わが国では、記録コードに対するJISがあるのみで、標準化の動きが弱いことは、データ交換に対するインパクトが弱いことを示しているが、これは需要が弱いことを意味してはいない。潜在的な需要が顕在的なものになるための前提条件として、チャネルが設定されていないことによるものであり、情報現象に直接関係している人達の間での、データ交換問題に対する態度が消極的であることに大きい原因があるともみべきであろう。

### (ii) 探索用ファイル

次に探索用ファイルについては、いわゆるIR(Information Retrieval)の初期において用いられたシーケンシャル・メモリーによるシリアル・ファイル時期を脱し、ランダム・メモリーの実用期となった。探索用ファイルに要求されるファクターは、

- ファイル・サイズのコンパクトさ
- 更新に要する時間の短縮
- 質問当りの応答時間の短縮
- ファイル構造の改善、拡張に対する柔軟性

○プログラミングの容易さ

○質問頻度（特にピーク時）に対する措置（プライオリティ、待ち等）

○ハードウェア構成のコンパクトさ

などである。これらを充たすために、ファイル技術として、いくつかの工夫がなされている。

1. ダイレクト・ファイル構造のレコードをシーケンシャルに配置したもの。

これは最も単純であり、更新、プログラミング、ハードウェア構成においては有利であるが、応答時間は長く、小規模のファイル以外の実用的でない。

2. ファイルをインデックス部分とデータ部分に分け、データ部分をダイレクト・ファイル構造とし、インデックス部分をタグとデータ部分のアドレスから成るインバーテッド・ファイルで構成するもの。

これは応答時間を可成り短縮できる。しかし、更新においてインデックス部分の内容を全面的に造り変える必要がある。

3. インデックス部分をリスト構造とするもの。

タグ記号が樹型表現できる場合には有効である。更新も、ポインターのおき方の工夫によっては容易にできる。

4. 相関係数によるファイル分割法。

データ集合の中から、あらゆるデータの対 (pair) をとって相関係数を求め、相関の高いもの同志を1つのグループにまとめる。このとき、これによって得られる各部分集合のものデータの数を大体等しくなるように分割する。各部分集合の図心値を求め、それをその部分集合の呼び出しコードとする。探索のときはクエリー・データの値と最も近い図心値を持つ部分集合を呼び出し、その部分集合内のみをサーチする。これは、ファイルが相当大きくても、サーチ対象を小さくできる利点があるが、データの追加、削除の激しいときは、部分集合（特にこの図心値）が安定しないため、更新がむづかしい。

これらの種々のファイル技術についての説明や効率比較についてはこの章末の文献を参照されたい。

ファイルの中でこれらの各レコードをメモリーの中に位置づける方法としては、

1. アイテム・キーの配列規則（番号順、アルファベット順、またはマニュアルによって指定された順など）を設定しておき、その順にシーケンシャルに配置する。

2. 前述と同じように配列規則によって配列されたアイテム・キーに対し、レコードの格納番地（およびレコード長、文字セットなど）を示すディレクトリーによって配列位置を明らかにする。

3. アイテム・キーを変換公式（中央二乗法、折りたたみ法、除算法など）によって格納番

地に変換する。このとき、すべてのアイテム・キーに異なる番地が与えられるようにする。

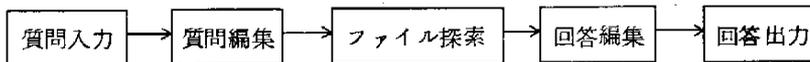
4. アイテム・キーに対して同様に交換公式によって格納番地を与えるが、このとき、いくつかのアイテム・キーに対して同じ番地を与えることを許す。このときは、同じ番地を持つアイテム・キー（これをシノニムという）の中の1つを正規の番地に格納し、他は別の番地に格納するが、その処理には、

- a 各アイテムにチェイン・フィールドをとり、次のシノニムの番地をチェインする。
- b シノニムを続く番地に格納する。
- c シノニムをすべて1つのバケットの中に格納し、全部を主メモリーに読み込んでサーチする。

などの方法がある。

## (2) 探 索

ファイルに対して質問を与え、求める回答を得るには、次のプロセスが必要である。



この中で、中心的な位置にあるファイル探索から現状を概観してみよう。

### (i) ファイル探索の技術

ファイル探索には、3つの部分がある。

#### 1. サーチ

後述する質問編集プロセスで作成されたクエリー・データに対し、求めるレコードが、(1)あるか否か、(2)あればどこにあるかを知るステップ

#### 2. 照 合

クエリー・データのヴァリューと、タグのバリューとを比較し参照して一致を求めるステップ

#### 3. クエリー・データの指定する構造を求めるデータが充分か否かの判定

#### A サ ー チ

サーチには次の方法がある。

##### 1. シーケンシャル・サーチ

シーケンシャルに配置されたレコードを、1つ1つサーチする。最も単純であるが、サーチ時間も最も長い。

## 2. 2分探索法

アイテム・キーの配列順序が定まっている場合、全アイテム・キーの中央位置を求め、クエリー・データのヴァリュアが中央位置より大きい半分に属するか小さい半分に属するかを判定し、更に該当する部分の中央位置との比較を繰返して求めるアイテムに到達する。シーケンシャル・サーチに比して著しくサーチ時間を短縮できる。

レコード数	シーケンシャル・サーチ		2分探索法	
	最高回数	平均回数	最高回数	平均回数
10	10	5	4	3
100	100	50	7	6
1,000	1,000	500	10	9
10,000	10,000	5,000	14	13
100,000	100,000	50,000	17	16
1,000,000	1,000,000	500,000	20	19

## 3. ダイレクト・サーチ

クエリー・データの値をそのまま求めるレコードの格納番地としてサーチする。これはファイルの構造または質問編集においてそれが可能なように準備しておく必要がある。

## 4. インダイレクト・サーチ

クエリー・データの値を変換公式（ファイル構成時に用いたものと同じ公式）によって求める番地に変換してサーチする。このときシノニムがあるときは、そのシノニム処理に合わせたサーチ手順を工夫する必要がある。

## B 照 合

ヴァリュアとヴァリュアの一致を見るのに次の2つの方法がある。

1. 完全一致法：ヴァリュアを構成する文字、数値、ビットが、すべて完全に対応したとき一致と判定する。

2. 部分一致法：クエリー・データのヴァリュアが、ファイルの中のデータのヴァリュアの一部分と対応がついたとき、一致と判定する。

完全一致法のみで判定する場合は、プログラムの簡便である。しかし、数値の場合「頭何桁まで一致すればよい」とか、文字の場合「語幹のみ一致すればよい」とか、センテレスをそのままファイルに収めたとき、「特定の地名、または人名、物理名等が含まれていればよい」という判定をしたい場合には、部分一致法を必要とする。この場合、次の3種がある。

1. 前方一致：「頭何桁の文字が一致すればよい。」という場合
2. 後方一致：「データの後何桁の文字が一致すればよい。」という場合
3. 中間一致：「データの最初でも最後でも、中間でも、指定する文字列があればよい。」

という場合。この場合、あまり単純で短い文字を指定するとノイズが起る。例えば、集積回路に関するデータを求めるために“IC”，同位元素に関するデータを求めるために“ION”などを中間一致で照合すると，romanticやcommunicationなどまで回答されることとなる。

この部分一致において、一致の始まるべき桁数を指定することも可能である。

### C 構造適合性の判断

今質問が、

日本と米国と西独の綿製品と化繊製品について1950年から60年までの製産高と輸出高を統計資料別に、資料名と数値とをリスト・アップしてほしい。ただし、1955年のみは不要である。

というものであったとする。そしてファイルの中には、収録項目のそれぞれ異なる統計資料が入っており、中には日本だけのものもあり、米国と西独のものもあり、綿製品だけの統計もあり、綿化繊の統計もあるとしよう。このファイルに対して、この質問は、入力前に次のように表現される。

「“対象国アイテム”のヴァリューは“米国”または“西独”または“日本”であり、かつ、“製品各アイテム”のヴァリューは“綿製品”または“化繊製品”であり、かつ“年次アイテム”のヴァリューは、1950より大きく、かつ、1960より小さく、かつ、1955ではない」という条件をみたすレコードのヴァリューが“綿製品”である“製品アイテム”に続く“生産高アイテム”のヴァリューと“輸出高アイテム”のヴァリュー、および、そのヴァリューが“化繊製品”であるような“製品名アイテム”に続く“生産高アイテム”と“輸出高アイテム”のヴァリューおよび資料名アイテム”のヴァリューを“資料名アイテム”のヴァリューでソートして出力せよ

この「条件」に合うか否かが「構造適合性」である。この判定には、次のような方法がある。

#### 1. 論理検索(ブーリアン・ロジックによる集合算)

上の質問の場合を単純に次のように表現しよう。

(AまたはBまたはC)かつ(DまたはE)かつ(Fより大きく、かつGより小さい)  
かつ(Hではない)

このアンダーラインの部分の部分を次のように記号で表現する。

または : + (or)  
 かつ : · (and)  
 ではない : ~ (not)  
 より大きい: ≧  
 より小さい: ≦

これにより、質問の構造判定条件は

$$“(A+B+C) \cdot (D+E) \cdot \geq F \cdot \leq G \cdot \sim H”$$

という論理式の形をとる。これは連言標準形であるから、選言標準形になおすと、

$$(A \cdot D \cdot \geq F \cdot \leq G \cdot \sim H) + (A \cdot E \cdot \geq F \cdot \leq G \cdot \sim H) + (B \cdot D \cdot \geq F \cdot \leq G \cdot \sim H) + (B \cdot E \cdot \geq F \cdot \leq G \cdot \sim H) + (C \cdot D \cdot \geq F \cdot \leq G \cdot \sim H) + (C \cdot E \cdot \geq F \cdot \leq G \cdot \sim H)$$

となる。

判定は各データのレコード内容が、この構造と一致したとき“1”とし一致しなかったとき“0”という2つの値をとると考え、この値に対してブーリアン代数を適用する。

$$\begin{array}{ll}
 1+1=1 & : \quad 1 \cdot 1=1 \\
 1+0=1 & : \quad 1 \cdot 0=0 \\
 0+1=1 & : \quad 0 \cdot 1=0 \\
 0+0=0 & : \quad 0 \cdot 0=0
 \end{array}$$

この結果が“1”となるとき、そのレコードを回答編集へ転送する。

これはL(A, B)を“1”と“0”の値をもつ論理関数とするとき結果において、

L(A · B) = 1 : ファイルの中のAで表わされる属性の値をもつ個体の集合とBで表わされる属性の値をもつ個体の集合の共通部分集合をとり出すことを示している。

L(A + B) = 1 : 同じくそれぞれAとBとで表わされる属性の値をもつ個体の集合同志の和集合をとり出すことを示している。

L(~A) = 1 : Aで表わされる属性の値をもつ個体の集合の補集合をとり出すことを示している。

L( $\begin{matrix} \geq \\ \leq \end{matrix}$ A) = 1 : Aで表わされる属性の値よりも大きい(小さい)値をもつ個体の集合をとり出すことを示している。

## 2. 重みづけ検索

コンピュータ処理において論理値計算よりも、加減算の方が容易であるため、論理検索を代用

する目的で用いられたり、単なるブーリアン代数でなく、質問項目に重みをつけて回答を制限したり、重みの高い順に回答を配列して利用を便利にする目的で用いられる。

原理的には、質問式の各アイテムに適当な数値を重みとして与えておき、ファイル内レコードが一致した項目に対し、and, or, not, >, <のいづれであってもすでに重みを加算する。そして、不必要なものを除外するために、加算した結果に対し、あらかじめ下限値を設定しておき、それ以上の値を1つレコードのみを回答と編収に送る。回答編集では、重みの高いものから順に回答プリントする。下限を設定しなければ、重み順に出力されたものを利用者が、この利用時に重みを参考としながら用いる。

### 3. 相関係数とする親近率判定

原理的には、クエリー・データとファイル内データの相関係数を計算し、その値がしきい値以上あるものを回答としてとり出す方法である。まだ実験的段階を出ないが、SaltonのSMARTという完全自動化を意図した検索システムの中で実用可能な手法にまとめられている。

#### (ii) 質問編集

質問が与えられたとき、その質問を入力したあと、ファイル探索に適した形にまとめる過程が必要である。それは次のプロセスを含んでいる。

1. 入力した質問のエラー・チェックとブルーフ・リスト作成
2. 質問式の構成
3. 質問式の拡張・制限
4. 出力条件の指定

(1)は一応常識的なことであるから省略する。

(2)の質問式の構成は、次のような要素プロセスの組み合わせである。

1. 質問の中のタグ・コードの桁数を数えて、レジスタに入れる。

これは照合過程が必要である。

2. 入力質問の論理式を変形する。

ファイル探索において標準形に直したり、 $\sim(A+B)$ をド・モルガンの定理を用いて $\sim A \cdot \sim B$ に変形した方が都合のよいとき必要である。

3. タグ・コードを、ファイル内レコードの配列規則に合わせてソートする。

サーチ手順を単純化してサーチ時間を短縮するために必要

4. タグ・コードを、ファイル内に出現する頻度(これはファイル構成時に、タリー・リストとして常に統計をとっておく)順にソートする。

質問式が連言標準形のときは、頻度の最も小さいものから探索を開始するのがよく、

選言標準形のときは、頻度の大きいものから探策を開始するのがよい。

#### 5. 予備探索用質問式を作成する。

探索速度を短くするため、予備探策によって該当レコードの範囲を狭めてから本探索を行なう方法をとる場合、前述のタグ・コードのファイル内使用頻度により、予備探索用質問式を作成する。例えば

○質問式

$$(A \cdot B) + (C \cdot D)$$

○各タグ・コードのファイル内使用頻度

A : 100

B : 20,000

C : 5,000

D : 500

とすると

予備探策用質問式

$$A + D$$

#### 6. 質問式のリスト表現

探策時間を短縮するため、質問式をリスト構造に表現する。例えば

○質問式：

$$A \cdot (B + C + D) \cdot (E + F)$$

○各アイテムの格納番地：

A : 100

B : 200

C : 300

D : 400

E : 500

F : 600

○リスト構造

Aから順番にファイル内レコードをサーチし、照合するに当たって、一致をみたときと不一致のときの両方について、次にサーチすべき番地をリストで表現する。

サーチするコード		一致したとき	不一致のとき
A	100	200	Reject
B	200	500	300
C	300	500	400
D	400	500	Reject
E	500	Hit	600
F	600	Hit	Reject

Hit はそのレコードと回答編集に移すことを意味し、Reject はそのレコードが回答編集に移されないことを示す。

質問式の拡張・制限は、シソーラスや用語管理システムなどの援助によって、より広い意味をもつタグ・コードや、関係の深いタグ・コード、要求する数値に許容範囲がある場合にその範囲に質問を拡張したり制限したりする必要があるときの手続きである。

- (1) 代入  $A \cdot B \rightarrow A' \cdot B$
- (2) 追加  $A \cdot B \rightarrow (A + C + D) \cdot B$
- (3) 拡張  $A \rightarrow > B \cdot < C$

などの組み合わせせる。このとき、シソーラスや用語管理のシステムとの、データの受け渡しやデータ処理のインター・フェースの充実が必要である。

出力条件の指定は、回答編集の方法がいくつかサブ・システムとして組み込まれているとき、そのサブ・システムを指定するものである。

#### (III) 回答編集

同時に何人かの利用者から、それぞれいくつかの質問を受けつける場合、ファイル探索においてはサーチ照合の時間を短縮するため、同じような質問はまとめて行なうし、その順番も必要に応じて組み変えてしまう。そのときは回答のために探索出力を、利用者毎にソートし、質問番号順にソートする必要がある。また重み順に回答結果をソートする必要もある。回答書のフォーマットが幾種類あるときは、指定されたフォーマットに合わせた出力編集を行なうことも必要である。更に探索によって得られた結果に対し、指定された演算(合計、平均値のように単純なものから、組み込まれた複雑な計算のサブ・プログラムを呼び出して結果を求めるような場合もある。)を行ない、その結果を編集することも含まれる。その出力がLPのみでなく、漢字プリンターとか、マイクロ・フィルム出力(COM)のような場合には、それぞれの管理システムヘデータを受け渡すための編集であることもある。

データの出力において加工度が高まるにつれて、回答編集のシステムは多くのサブ・システムの複合となり、探索システム全体の管理ソフトウェアを必要とするようになる。

このデータ加工において、ファイル側が守らねばならぬ原則がある。それは

加工はユーザの要求によって行なうもので、ユーザの選択を保証しなければならない。ことである。ファイル管理者がユーザの要求によらずその意志によって加工を行なうことは、情報管制につながるものであり、情報全体主義による社会のコントロールを生む危険があるからである。

### (3) プログラム言語

ファイルの構成や探索のアルゴリズムはアルファ・ニューメリック・データに関して、現状では概略を述べて来た考え方の組み合わせで表現できる。そのプログラミングに当っては、各機種毎のアセンブラによるものと、COBOL, PL/1, FSRTRAN, などが用いられる。アセンブラによるものは、微視的効率はたしかに良いが、システム全体が発展期にある場合は、巨視的に見てコンパイラによる方が有利である。また、データをプログラム・パッケージをつけて供給する場合、コンパイラの必要がある。

しかし、現実的問題として、アルゴリズム自身はあまり変わらないのに、ハードウェアの発展やデータ量の増加と共に機種のスケール・アップを必要としたとき、ソフトウェアの全面的書き直しを必要とする場合がある。また、汎用言語で書かれていても、例えば供給をうけたものがCOBOL/65で、自社で用いるのがCOBOL/61であったりすることもある。

このような問題から

1. 基本的なアルゴリズムは大体設定されているのであるから、それを最も扱いやすいファイル処理用のコンパイラの開発
2. 基本アルゴリズムから求めるプログラムを造り出すゼネレータの開発

が実際問題として要求される。COBOLは、データ・ディビジョンの繁雑さ、文字ストリングの処理やリスト構造の処理などに難点があり、SNOBOL, COMITは文字ストリングの処理はよいが他の機能が弱く、LISPはリスト構造にはよいが他の機能がな。など、現在のコンパイラはファイル処理に難点がある。公共的な技術開発のテーマとしてとり上げるべきものの1つである。

#### (4) オンライン・システムおよびマルチ・ベースの管理問題

日本では、バッチ処理の探索プログラムは実用段階に入り、コンピュータ・メーカーも大体、その機種に合ったプログラム・パッケージを持っている。しかし、オンラインの探索プログラムはまだ開発段階とみることができる。この点、米国では実用段階に入っていることからみて、数年の遅れがあるとみられる。

ファイルの探索をバッチ処理で、しかも1回の処理で十分な回答を得ることは、ファイルの構造や、質問の内容が複雑になると、実際問題としてむづかしい。質問を分析し、探索式にする過程で、よほど熟達した人間を必要とする。しかし、この人材の養成は、一般情報処理技術者訓練の中でも、むづかしい部分である。

このために、ある程度の基礎訓練をうければ、システムの支持をうけて、試行錯誤的に正しい回答が得られるように、ファイル処理のシステムを造ることが望ましく、究極的には、ファイルの利用者が、自分でファイルに直接アクセスできることが望ましい。しかし、

1. ハード的制約（探索時間、ファイル・サイズ）
2. 通信回線の制約（容量、速度、使用料）

からみて、すべてバッチ探索に代わるものではなく、実用可能性には限界があるとみるべきである。その条件は、

1. ファイル・サイズが比較的小さいこと。
2. レスポンス時間あたりのコスト・パフォーマンスの良いこと。
3. 試行錯誤過程が有効であること。

この点からみて、

○後述（§7のネットワーク）する倉庫型データバンク

○有線オンラインによる組織内システム（通称インライン）

○質問式確定までを探索や会話形式で行ない、簡単なメッセージで表現できる内容のみをオンラインで回答し、本格的なファイル内容の詳しい出力をオフラインで別処理するかバッチ処理に回わし、詳細な回答をあとで送るといふ併用システム

などが実用的である。

##### (i) オンライン・システムの構成

オンラインでのファイル処理は、会話形式を処理するメイン・プログラムと、それによって呼び出されるいくつかのサブ・プログラムから成る。メイン・プログラムは、いくつかのタスクの集合と、その全体の管理プログラムから成る。（第1.6 図参照）

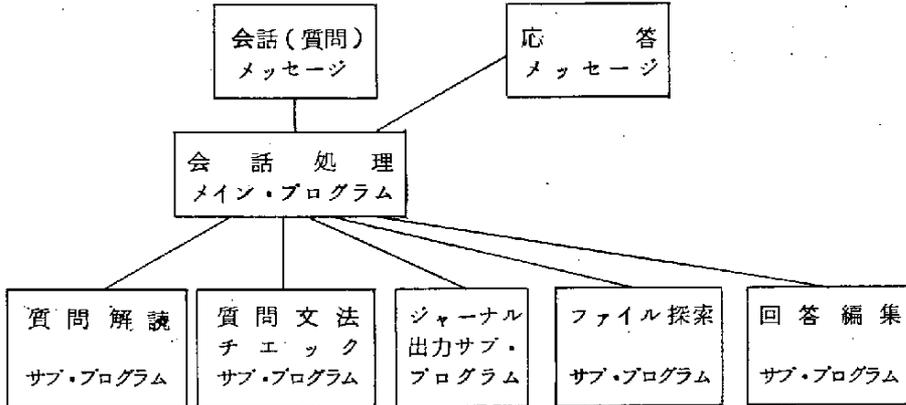
メイン・プログラムを構成するタスクは、例えば次のようなものである。

1. ライン・コントロール

端末から適宜入って来る文字を、メッセージに組み立てる

2. インプットメッセージ処理

このメッセージを内部コードに変換し、次にどのタスクに送るかを決定する。



第 1.6 図 J I C S T の実験的オンライン探索システムの構成

3. 待ち行列コントロール

次から次と同じタスクにメッセージが入って来たとき、その各タスク別にできる待ち行列をコントロールする。このとき優先順位も処理する。

4. 会話開始処理

入力した、利用者コード、仕事の種類、等をチェックし、探索質問や回答編集指示などを受付ける準備をする。

5. 質問文処理

入力した質問の文法チェックをそのサブ・プログラムを呼び出して行なうための処理

6. ファイル探索

質問文を解説し、探索指令を作成し、ファイル探索を行なうため、それぞれのサブ・プログラムを呼び出して実行させ、結果をドラムなどに格納する処理

7. 回答要求処理

回答要求のメッセージを文法チェックし、探索タスクへ送る処理

8. 回答出力処理

回答を編集し出力メッセージを造る処理

## 9. 出力メッセージ処理

メッセージを出力コードに変換し、それぞれの端末に送るための処理

## 10. 時間監視処理

これはあくまで例示であり、システムによって、それぞれの工夫がこらされている。(第1.7 図参照)

### (iii) 会話の形式

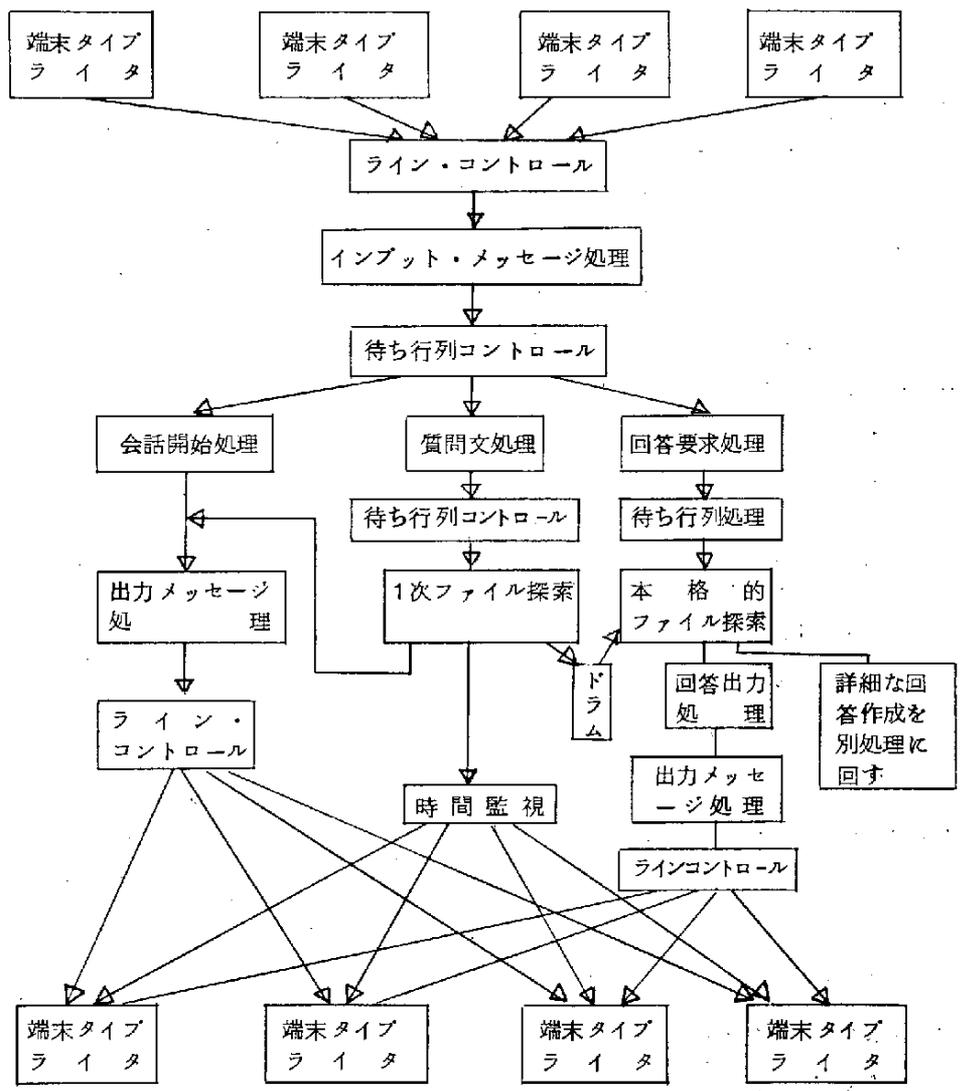
会話は、システムによって設定されているいくつかのコマンドのタイプ・インと、システムからの回答によって行われる。コマンドの数と種類は、システムによって異なるが、大別すれば、

1. 会話の開始や終了に関するもの
2. 探索に関するもの  
: 質問式の入力、その拡張、定義に関するものなど
3. 回答要求に関するもの  
回答の種類の手指示、出力形式の手指示など
4. 補助的コマンド

利用者が質問式決定のために必要な情報を出力したり、会話進行中の途中の結果を記憶させ、必要に応じて出力させたりするのに用いる。

この会話の進行過程の例は第1.8 図を参照されたい。

第 1.7 図 J I C S T の実験的オンライン探索システムのタスク構成



第1.8図 JISTの実験はオンライン探索システムの会話例

REQUEST

DOORS:

" JICST オンライン IR システム 71-02-01 11:18:26 "

" アナタノ コードヲ イレテクダサイ "

G393:

" ジュンビ ワ オワリマシタ "

" コマンドヲ イレテクダサイ "

SEARCH A01=LISP+SNOBOL:

" タンサクヲ ハジメマシタ "

" シバラク オマチクダサイ "

" A01 ワ 13 ケンデス "

" コマンドヲ イレテクダサイ "

SEARCH B02=GIS:

" タンサクヲ ハジメマシタ "

" B02 ワ 28 ケンデス "

" コマンドヲ イレテクダサイ "

SEARCH C03=MARK-4:

" タンサクヲ ハジメマシタ "

" C03 ワ 11 ケンデス "

" コマンドヲ イレテクダサイ "

DEFINC D04=A01\*B02

" D04 ワ 7 ケンデス "

" コマンドヲ イレテクダサイ "

HELP:

ユーザがコマンドとして何が使えるかを知りたいとき、説明つきで全コマンドを出力する。

" コマンドヲ イレテクダサイ "

OUTPUT D04, 1, 3:

" 80300653 80700735 80700047 "

" コマンドヲ イレテクダサイ "

MORE 2:

"80600358 80900175"

"コマンド ラ イレテクダサイ"

PRINT A01, 6, 13;

"コマンド ラ イレテクダサイ"

SAVE A01;

"コマンド ラ イレテクダサイ"

END:

"ベツノ シツモン ラ ツヅケマスカ?"

NO:

"サヨナラ 71-02-01 11:28:26" 10MOOS,

03M4.7S077MS

## V センテンシアル・データおよび アナログ・データの処理技術

アルファニューメリック・データについての技術を中心に、ファイル処理技術の現状を紹介した。これを見てもわかるように、センテンシアルな情報も、現在の所、「文字ストリング」とみなして（ソースの援助はうけるとしても、「文字ストリングの照合」により処理している。これは「文の意味」による処理ではない。

センテンシアル・データの処理は、究極的には、「その意味する内容が質問の意味する内容をみたすか？」のレベルに到着しなければならない。しかし、そのレベルに到達する道は、長いのである。

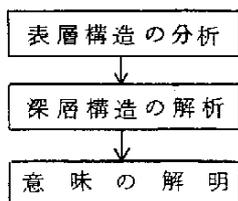
### (1) 文法理論、機械翻訳

文法理論は、若き天才といわれる N. チョムスキーの提出した革命的な文法論により、1950 年以後、急速な発展を見せた。これは、

1. 文法を、生成規則と変換規則の有限な集合とみなし、演繹の体系にまとめる。
2. 人間（でなくとも言語を用いる主体）が、自己の表現したい内容を「深層構造」の表記法の中で表現すると考える。そのときの規則が生成規則でこの段階を句構造レベルと呼ぶ
3. 深層構造の表記法で与えられたものを日常のコミュニケーションで用いられる表層構造の表記法に変換する。このときの規則が変換規則で、この段階を変換構造レベルと呼ぶ。
4. 意味の問題（Semantics の領域）は、この文法的シンタックスの解明の次のステップとして与える。

という考え方である。これを一般に「変換文法理論」と呼ぶ。

センテンシアル・データの処理は、これの逆ステップを踏み



という形をとる。この表層構造の分析は「構文解析」と呼ばれ、機械翻訳のプログラムの中では何らかの形で行われている。しかし、現在の構文解析法では文法的には正しくても、意味的には

ナンセンスな解析結果や、1つの文が幾通りもの意味に対応する多くの構造に解析できることによる。Syntactic Ambiguityを避けられない。これが機械翻訳の実用化を妨げている。

この構文解析技術は、欧州系の国語については、実用レベルの解析システムがいくつか発表されている。しかし、日本語については、日本語の文法がチョムスキー流の変換文法的にまとめられていないため、実用レベルの構文解析システムがまだ生まれていない。

オーソドックスな変換文法が

$$S \rightarrow NP + VP$$

の形の、「書き替え規則」の集合であるのに対し、構文解析や、語のストリングのパーシングに便利のように、集合論的文法やネットワーク型文法なども発表されている。

表層構造から深層構造への逆変換についても、種々の研究が進められている。マッコーレイやウッズ達は、述論論理式の構造を深層構造の表規法に用いる可能性を検討しそれをウッズは簡単なモデルで質問応答システムに利用している。

日本では、機械翻訳の研究から出発して、電子工学者の方からの言語への接近、それに刺戟をうけた言語学者から言語の機械的処理問題への接近がみられるが、それらの部分的研究を統合するプロジェクトや十分な予算の支出がないため、研究の進み方が欧米、特に米国に比して遅れを示している。日本での代表的なものをあげると、

電子技術総合研究所	情報基礎研究室
京都大学工学部	坂井研究室
九州大学工学部	栗原研究室
国立国語研究所	
東京女子大学	水谷研究室
日本科学技術情報センター	資料部

などがあり、意見交換の場として、情報処理学会の Computational Linguistics 委員会がある。

機械翻訳は、世界でもまだ実験段階といえる。日本にも、電子技術総合研究所、京都大学坂井研究室などに実験システムがある。最近、原子力関係の国際機関、EURATOMで、翻訳サービスに実用している例があるが、まだ実用試験とみなすべきで、普及の段階ではない。

## (2) 自動分類，自動索引，自動抄録

ファイルに構成の中の、分析・評価の手続きを自動化しようとする試みは、ルーンの自動抄録の試みを出発点として、主として米国で精力的に行なわれた。この系統の米国での努力は、機械

翻訳が言語理論を基礎とするのに対し、主として統計理論を用いた。

ルーンの試みは“a”“of”“the”等の意味そのものを担わない語を除き、語が文献の中で使用される頻度から「重要語」をとり出し、それを密度多く含む文を「重要な文」とするという基本的考え方に立っている。

この種の考え方が、自動分類や自動索引にも導入され、同じ標題に表われる語の共出現頻度によって語の関連係数を求める考え方も生まれた。更に因子分析法の導入、数量化理論の導入など多くの手法を用いて、この種の研究が進められている。

しかし、まだ実験的研究の域を出ず、言語現象を確率過程とみる理論形成も、また実用に足る成果も見出されていない。これは、人間が、内容的に似たものをグループしたり、相違点を識別する働きは、図形や音声の認識や識別よりもはるかに高度な働きであり、「思考パターン」の分析と等価であることによる。

日本で、現在この種の研究を行なっているのは電子技術総合研究所、日本科学技術情報センターなどがあるが、十分な研究資金や研究環境がないため全般的には低調である。

### (3) 質問応答システム，事項検索

MISなどで、記録文書や管理文書や契約書の龐大なファイルの中で、今議論している問題に直接関係のある記述を、そのままとり出したり、何らかの質問を行なうと、ファイルの中にあるいくつかのレコードを組み合わせて推論を行ない、その答を回答して来るシステムは、潜在的要求として相当高いものである。

ファイルの中から関係のある記述そのものをとり出すのを事項探索と呼び、推論過程を経て答えを出すシステムを、質問応答システムと呼ぶ。

これに対して、米国では、文の構文分析技術の発展の1つとして、地道な研究があり、日本では日本科学技術情報センターに、基礎的な研究がある。

米国の例をあげると、ハーバード大学のウッズが“Air port guide”という実験場面をとり、

「フライト#JL005の飛行機は、羽田を、1971/7/1/10:00に、サンフランシスコへ発ち、1971/7/1/7:00に着く」

という文を、

LEAVE(JL005, HANEDA 1971/7/1/10:00,  
SANFRANCISCO 1971/7/1/7:00)

というように表わす。このような表記法で、種々の飛行機の発着、経由地、その日時を表現し、

ファイルに収める。この外、種々の空港内の案内情報、料金、サービス内容等を、ファイルに格納する。そして、予想される質問と、それに対する回答を考え、それを導びく推論アルゴリズムを用意する。

これに対し、「何月何日にA地に寄り、何月何日何時までにB地に行き、そこを何月何日の午前中に発って、同日の何時までにC地に着くには、どの経路で、それぞれ何の飛行機に乗ればよいか？」という質問を出すと、その可能性を調べて、必要情報をすべて出力して来るといふ実験システムである。

これは場面は単純であるが、

Predicate (Augment 1, Augment 2 ………)

という形に文を変換する考え方の中には、文の深層構造を、述語論理式の形で表現する基本的考え方があり、日本文(表層構造)のそれによる表記の妥当性を調べるモデルとして、簡単な実用可能なシステムで実験しているのである。

この程度のものであれば実用可能であり、日本でも、このような程度のシステムを実用するならば、その一応のシステム構成は可能である。しかし、この奥にある基礎研究において、日本文法の未整理であることが、より複雑なシステムの実現において障害となることは、残念ながら認めざるを得ない。

#### (4) アナログ・データの処理

アナログ・データの伝送においては、ファクシミリ、写真電送、TV電話などがある。又、電送技術において走査密度が上り、マイクロ・フィルムの電送も実験室的には完成している。この系統での問題点は、写真や全面が色彩や濃淡の階調のあるデータでないかぎり、文書や図面において、情報の乗っている、文字や線が黒く印字されている部分が少なく、白い部分が圧倒的に多い。この冗長度を減らして、伝送容量を増すことである。

アナログ・データのコンピュータ処理は、現在の所、一度デジタル情報に変換しなければならない。その常識的方法は、OCR、ライト・ペンであるが、今後の問題としては、可干渉光(レーザー光など)を利用した光学的情報処理技術とエレクトロニクスの結びつきである。

この中で、ホログラムの利用は魅力的である。特に、立体像の取扱い、3原色に分けた色彩像の取扱いが、フィルム上の小さいセル上での、感光・非感光のbinaryなデータに変換できる点で興味がある。ホログラムを利用して、手書き図形のコア上への入力、日本電気KKが発表しているが、それからの発展が期待される。特に、光学的情報処理技術を利用して、部分的図形情報からオリジナルの複元や、小さな模型からプレス原型の作図などをコンピュータ処理する技術

などは、アナログ・データのファイル格納やとり出しに応用可能である。

出力系統では、入力や処理と異なり、技術的条件が簡単になるため、多くの実用化が行なわれている。その1つはCOMといわれるマイクロ・フィルム出力である。それと並んで、日本語特有の、漢字カナという、多字種、高分解能を要する文字の処理である。

漢字穿孔鍵盤によりパンチで入力し、フィルム上の字母を電子工学的に走査してCRT上に出力するGEM3800システムは日本科学技術情報センターで最初に実用し、その後数カ所で実用している。これに対し、これよりは機能は少し落ちるが、高千穂交易社の漢字出力システムが発表されている。

米国でも漢字プリンターは研究されており、アナログ形のまま字母を入力すると、それをコア上に記憶し、入力原稿のデジタル信号によって字母をコアからとり出してCRTに出力するRCAのVIDEOCOMPが、漢字プリンターとして最も実用的である。その他、多くの実用システム、実験システムがある。例えば、光線又は電子線を量子光学的に又は電磁的に偏向させることにより、CRTや感光面に文字を画かせるシステムのような、特定の字種に限らない方法なども現われている。

アナログ・データのファイルと、直接探索するシステムとしては、フィルム上に、アナログ・データと、その探索タグを(黒白のbinaryコードで)写し込んでおき、タグでデータを探し見かけたものをすぐプリントするシステムがある。現在市販されているものでは、イーストマン・コダックのMIRACODE Systemが代表的である。これは、プリントされたハードコピーをユーザに手渡すか郵送しなければならない。

オフラインでは、種々の可能性がある。この中で実用的に問題となるのは、相当量の1次データをフィルム上に記録して、あらかじめユーザに送ってしまう。データを必要としたとき、その質問に対して探索を行ない、その該当レコードの、フィルム上アドレスを回答しユーザがそのレコードをフィルムからとり出す方法である。これは、ユーザの数が多く、フィルムの配布量が多くなると、超マイクロ・フィルムが有利である。これは富士写真フィルム社の超マイクロ・システム、NCRのPCMI(Photo-Cromic Micro-Image)が利用可能である。

## VI ネットワーク

ファイルを、情報の流通過程の中で考えるとき、ファイル単独では機能し得ない。

1. 収集端末および提供端末との結びつき
2. ファイル間の結びつき

が必要となる。この結びつきを以下でネットワークとして捉える。この際、通信回路によるオンライン・ネットワークのみでなく、データ交換用ファイルを経由するオフラインの結びつきも考慮する。このため、ここでのネットワークは、情報流通における機能的ネットワークである。

今、以下のものをノッド (node) とするとき、

データ・ソース	S
収集端末	CT
ファイル・システム	F
提供端末	ST
ユーザ	U

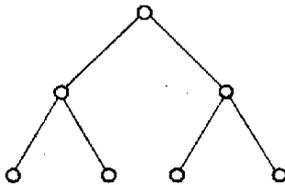
そして、各ノッドの間でデータの移動（オンラインであれ、データ交換用ファイルを通してであれ）があるとき、移動の方向に向けて矢印

A	→	B	ノッドAからBへの移動
A	↔	B	ノッドAとBの間での相互移動

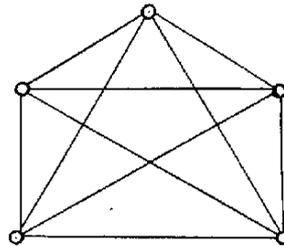
をつけると、ネットワークは方向性をもつグラフ（又はトランジション・ネットワーク）として表現される。（ただし、方向性がどちらでもあり得るときは、矢印をつけずに表現することとする。）

### (1) トリー（樹型）ネットワークとメッシュ（網目）ネットワーク

以後の議論において、ネットワークを樹型の構造をもつトリーと、網目の構造をもつメッシュの2つに類型化して扱おう。

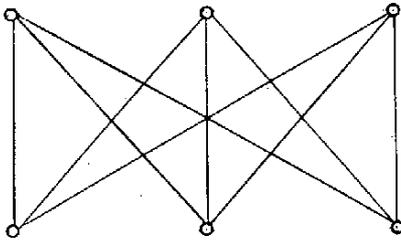


トリー



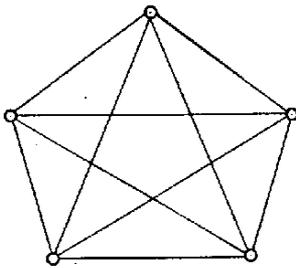
メッシュ

又、次の場合も複合トリーとして扱う。

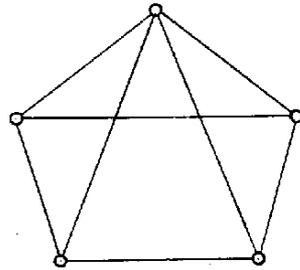


複合トリー

メッシュの場合、すべてのノッド間に連結のある場合を完全なメッシュ、それからいくつかの連結が取除かれた場合を不完全なメッシュと呼ぶ



完全なメッシュ



不完全なメッシュ

トリー・ネットワークの利点は

1. 構造が簡単であるため、設計・構成・維持が容易である。
2. 事故の発見、事故原因の究明が容易である。
3. 発信（データ供給）主導型のネットワークに最適である。
4. 多数のユーザからの要求が、シングル・データベースによって満たされる場合に有効である。

しかし、欠点としては

1. 「根」となるノッドと「葉」となるノッドの間には1本（中継ノッドを経由するとしても）のチャンネルしかなく、その容量を越えると通信不可能となる。
2. 同一チャンネル上にあるノッド間の通信は容易であるが、他のノッドとの通信には多くの中継ノッドを経るため、通信効率がよくない。
3. 「根」に近い所でチャンネルに事故があると広い区域に対し通信不能となる。

4. 1つのファイルが事故でつぶれた場合、致命的打撃をうける。

これに対し、メッシュの利点は、

1. あらゆるノッドが直接に通信できる。
2. どこかのチャネルが事故にあっても、オーバーフローしても、必ず他のチャネルを通しての通信が可能である。
3. ユーザが多角的な要求を持ち、シングル・データベースではみだし切れないとき有利である。
4. いくつかの、オブティマムなオーバーラップをもつファイル間をメッシュで結ぶことにより、1つのファイルが事故でつぶれた場合の危険を吸収できる。

これに対し欠点としては、

1. ネットワークの設計、構成、維持が複雑となる。
2. どこかのチャネルに事故が起っても、通信は保持できるため、モニタリングのシステムが強力でない、事故の発見・対策が遅れる危険性がある。

以上のように、トリーとメッシュは利点と欠点が相反しており、実際のネットワークはこの両者の併用となる。

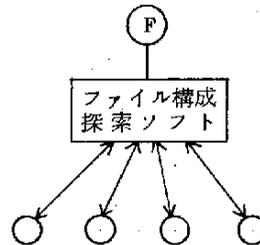
## (2) クラブ型ネットワーク

このクラブ型というのは、データ・ソースとユーザーが同じであるような集団の中でファイルが利用される場合をいう。ソースでありユーザーである集団が、ファイルを通してクラブを形成すると考える。

これは、従来のマス・コミュニケーションでは実現しなかった、メディアム・サイズのコミュニケーションという観点および、ファイルを通して形成されるコミュニティという観点から、今後の社会のあり方に示唆を与えるものとして興味のある場面である。

この場合、ファイル・システムにファイル構成および探索ソフトウェアがあり、利用者はデータ入力と質問入力することによって回答を得る。

この場合はファイル・システムと利用者間にトリー結合がある。この例としては、ヒューストン・クレジット・ビューロー、バンカー・レーモア社、バンク・オブ・アメリカ社などがあげられる。



### (3) 倉庫型ネットワーク

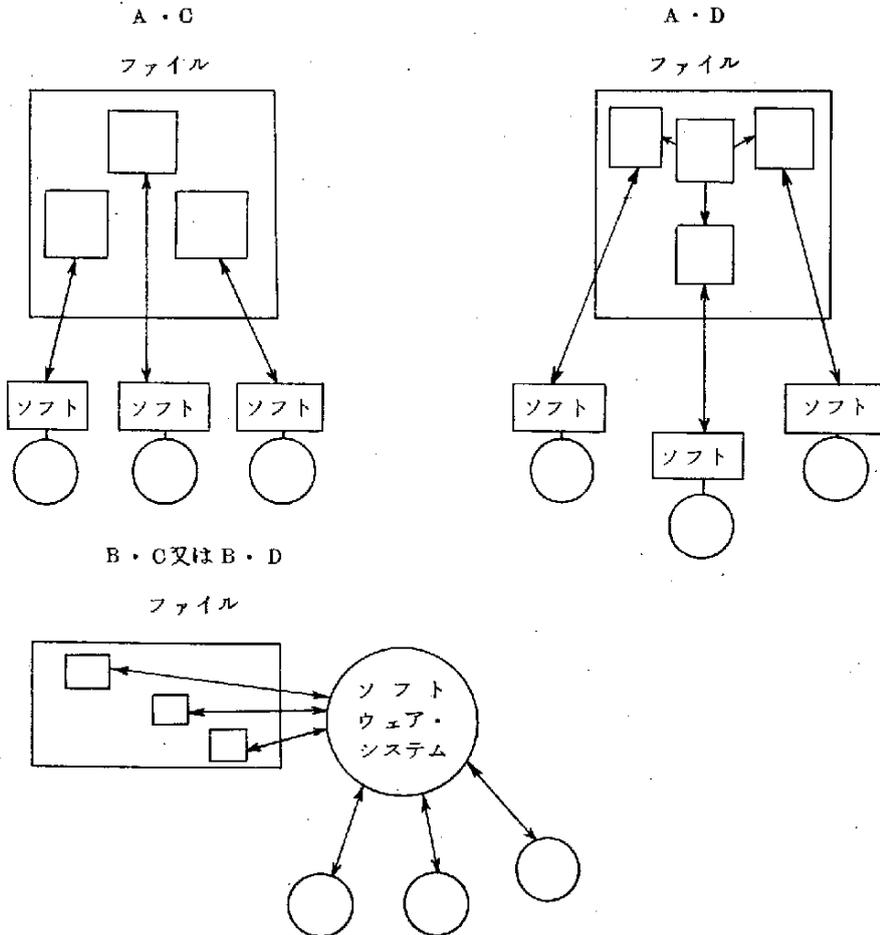
この場合は、利用者は自分のデータをファイルに預けておき、それを自分でとり出して利用する場合で、ファイルが倉庫業的な形態をもつ。このとき、利用ソフトウェアが、

- A 利用者自身が持つ場合
- B ファイル側で種々のソフトウェアを用意しておき、利用者が必要に応じてソフトウェアと呼び出して利用する場合

に分けられるし、またファイル自身が

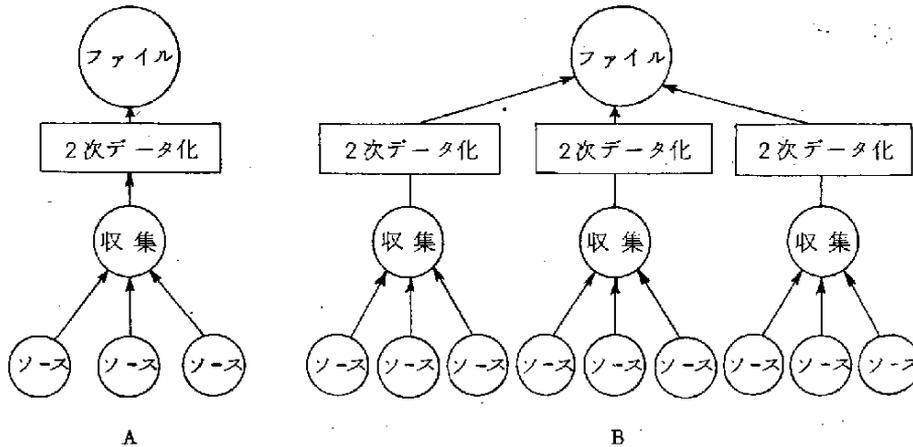
- C 利用者が自分で入力したデータしか利用できない場合
- D 多くの利用者の入力したデータを、ファイル間で交換し共同できる場合

が想定される。このB・Cの組合わせの例として、日本電電公社の科学計算サービス・システムがある。



#### (4) データ・ソースとファイルのネットワーク

データ・ソースとユーザは一般には異なるものである。そのとき、データ・ソースとファイルは、一般的には、トリー連結をする。このとき、収集端末と2次データ化機能の配置によって、いくつかの変形がある。



- A : ファイルに直結する収集端末と2次データ化機能にソースがトリー連結する場合。これの例は数多くあるがJICSTなどはその最も良い例である。
- B : 分類した収集および2次データ化機能にソースがトリー連結し、その収集、2次データ化機能がファイルとトリー連結する場合。Chemical Abstract Service (CAS) や原子力文献の国際的情報システムである International Nuclear Information System (INIS) などがこの例である。INIS の場合は各国に収集、2次データ化のセンターをおく計画でプロジェクトを進めている。CAS は現在の所は各国の個人単位の抄録作成協力者を収集および2次データ化端末としているが、現在 INIS 同様、各国に入力および出力 (すなわち情報サービス) の地域センター (欧州に2カ所、東南アジアに1カ所等) をおき、全世界を結ぶネットワークを形成する計画をもっている。

#### (5) ファイルとサービス端末とユーザの間のネットワーク

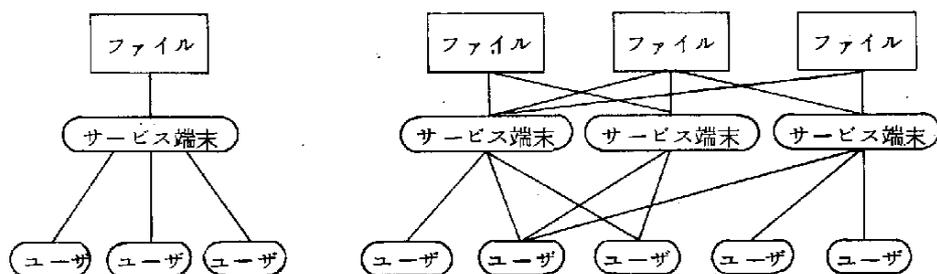
ファイルとサービス端末の間も一般にはトリー連結である。この場合、サービス端末が、サービス用ファイルを持つ場合と持たない場合で、技術的には大分異なる。

(i) サービス用ファイルを持たないサービス端末

この場合、サービス端末は

1. 要求の受付、ファイルへの伝送
2. 回答の受領、要求者への伝送

のみである。このファイル、ユーザとのネットワークは、トリーである。ファイル、サービス端末に分業が成立するときは、複合トリーとなる。



(ii) サービス用ファイルを持つサービス端末

これは、自分では収集、ファイル構成機能を持たずファイルの供給をうけて自分の所にサービス用のファイルを持ち、それによって自分の持つユーザにサービスする。これもファイルを持つ点からデータバンクの一種である。最近、マシン・リーダブル形式でのデータ供給が多くなっており、この種のサービスが日本でも数多く生まれる可能性がある。例えば紀伊国屋書店の科学技術文献に対するSDIサービスなどは、この最もよい例である。

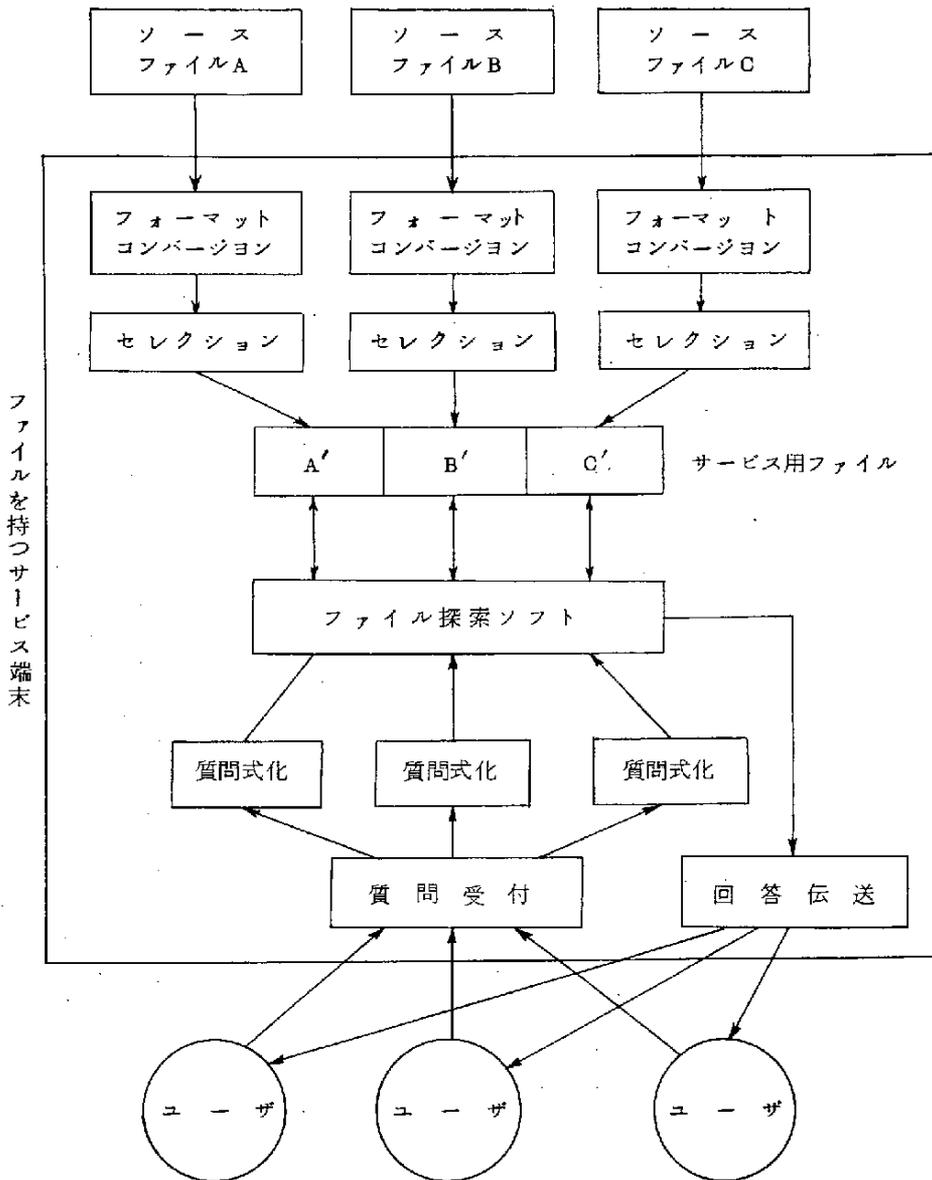
サービス用ファイルには、2つの考え方ができる。

1. 供給されるファイルがいくつかあり、その間でフォーマット、タグ・コードが標準化されていないとき、それらを共通のハードウェアと探索ソフトウェアで処理するため、フォーマット・コンバージョンを行なうが、タグ・コードのコンバージョンは行なわない。このときは、フォーマットは統一されているがファイルの内容は、いくつかのファイルの共存で成立っており、ユーザからの質問を、それぞれのサブ・ファイル向けの質問に分解し、出力結果をまとめる必要がある。

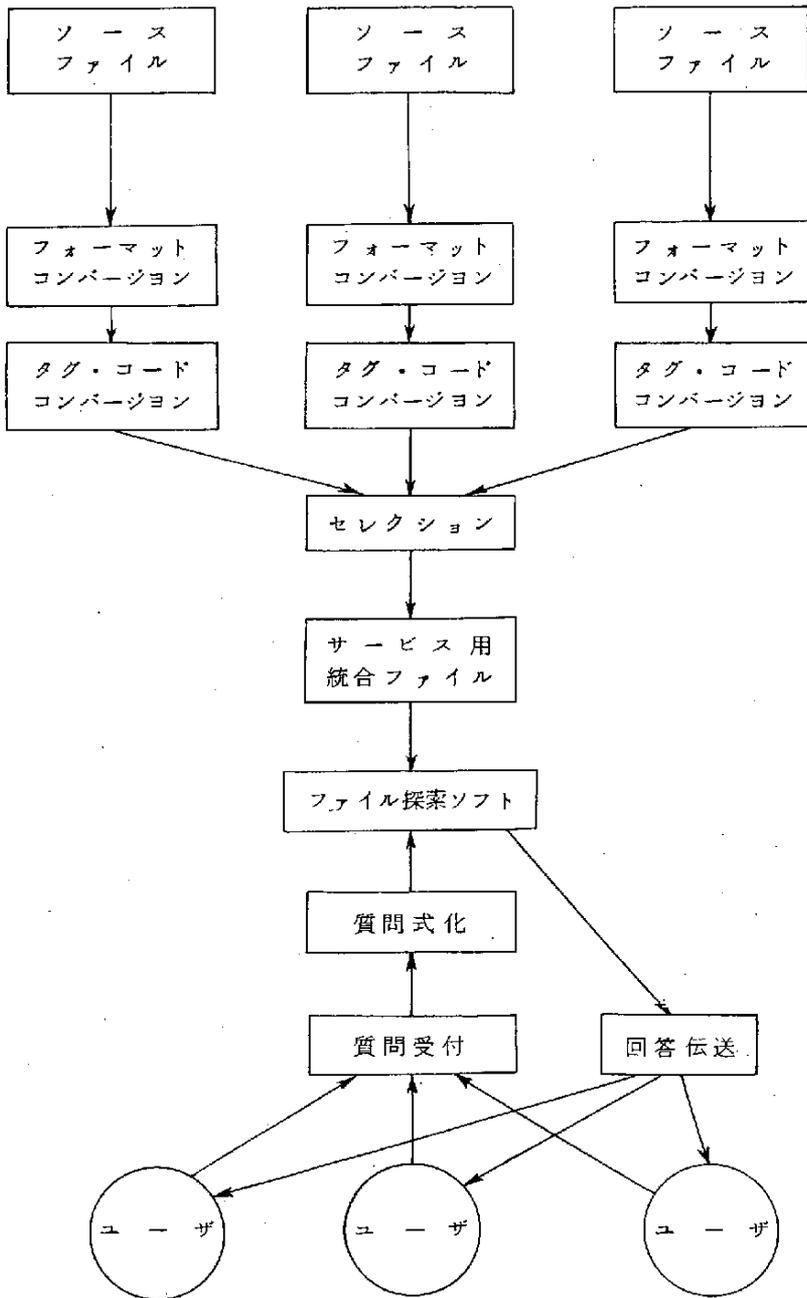
2. フォーマットのコンバージョンと共にタグ・コードのコンバージョンも行ない、供給された異種のファイルを、1つのファイルに統合する。このときは、タグ・コードのコンバージョンにあたり、ソースラスのコンバージョンを必要とする。ソースラスのコンバージョンの問題は今の所まだ実用レベルで解決されているとは言えない段階である。

(1)と(2)を比較すると、観念的には統合ファイルの方がすぐれていると思われる。しかしファイ

第 1.9 図 多種共存ファイル



第 1.10 図 統合されたサービス用ファイル



ル・サイズが大きくなり、サーチ時間が増加することを考えると、必ずしも有利でなく、特にオン・ライン・サービスの場合に実用不可能になる可能性がある。

共存ファイルを用いる例では紀伊国屋書店のSDIサービス、またJICSTの国際的な2次データMTの導入の最初のステップなどがある。この両者の比較検討をしている例では、米国の国立農業図書館(NAL)が、システム分析を行なっている。

## (6) ファイル間ネットワーク

あらゆる目的をみたくシングル・データ・ベースはあり得ない。商品の流通機構も、商品の種類によって、それぞれの流通機構があるように、情報の流通機構も、それぞれのファイルを通しての流通機構が必要となる。

しかし、ユーザの要求は多面的なものであり、1つの業務のために多種類のデータが必要となる。オペレーショナルな業務においては主として定型的なデータ利用となるため、固定したチャネルでよいが、プランニング業務においては、必要としたときに、どの種類のデータでも入手できることが保証されなければならない。

情報の流通機構が発展するにつれて、多くのファイルが、それぞれのシェアをもって社会の中に散在するようになる。このとき、ユーザは、自分の要求をみたくファイルがどれであり、どこにあるかを判定することが困難になる。商品の場合にも、百貨店、スーパー店などが生まれ、問屋街が生まれるように、ユーザはできるだけ手近かな、数少ない(できれば1つ)のファイルですべての要求が満たされることを希望する。

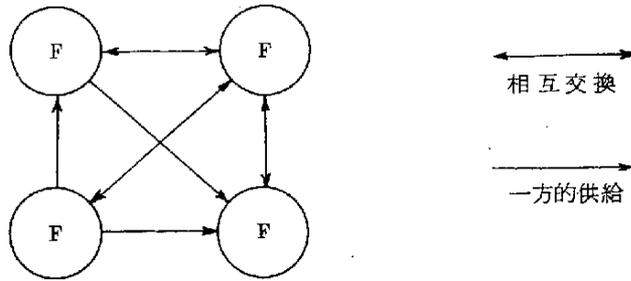
ファイルには分業があり、ユーザはできるだけ1つのファイルにのみアクセスして、すべての要求をみたそうとする。

という矛盾は、ファイル間ネットワークによって解決される。

ファイル間ネットワークは、トリーを基本に考えることは危険である。トリーはファイル間の自由なデータ交換を否定する考え方に立つものでファイルを系列化する思想が根底にあるからである。それは情報管制の危険をすら秘めているからである。選択の自由とユーザの平等の権利を前提とするファイル間ネットワークは、メッシュを基本とし、必要な場合はトリーも、不完全メッシュの一種と考える立場に立つべきである。

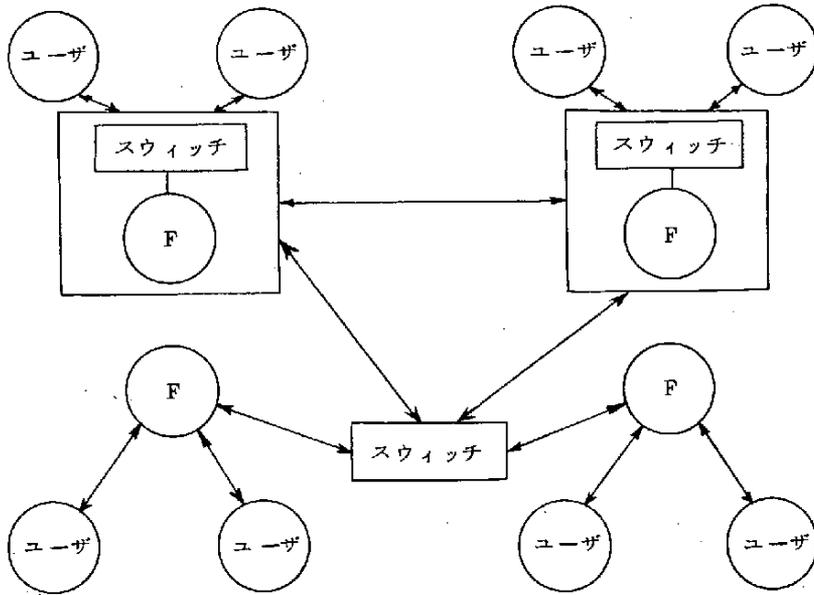
### (i) 定常的データ交換によって結ばれるファイル

ファイル間のデータ交換が、ファイル間の協定・契約により、定まったデータを定常的に交換する場合である。このときはスイッチ機構を必要としない。このとき、個々のファイルは、共存、統合の両方の場合があり得る。



(ii) 要求に応じてデータ交換が起る場合

ユーザの要求が、自己の持つファイルでみたされないとき、スイッチ・機構を通してそのデータを持つファイルを見出し、そのファイルにアクセスする。



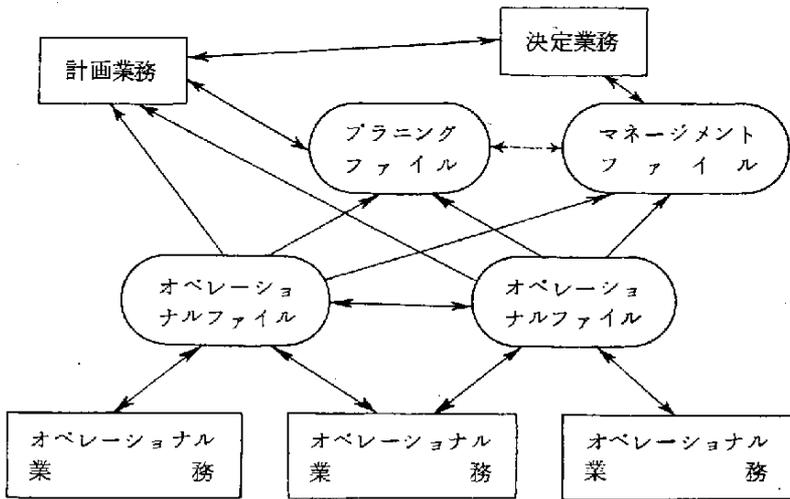
ファイルが箇々にスイッチ機構をもつ場合と、1つのスイッチ機構がいくつかのファイルを結ぶ場合があり、回答編集と伝達を要求を受付けたファイルが行なう場合と、質問の転送をうけたファイルが行なう場合がある。

スイッチ機構は、それ自身もまた「どのファイルに、どのデータがあるか」というファイルから成っている。このスイッチ・ファイルを個々のファイルが自から構成することは、情報の流通機構の発展と共に困難となり、スイッチ・ファイルの構成と供給という分業も必要となる。これが、インフォメーション・クリアリングハウスという名称で呼ばれるものであり、情報産業の発展期において、スイッチ・ファイルの供給もパブリック・セクターの果たすべき役割の一つである。

### ④ ファイル間に階層構造のある場合

これは組織内データ・ベースのあり方の1つである。オペレーショナル業務のためのファイルがいくつかあり、その中からマネージメントのために必要なデータをまとめた。ファイルや、プランニング業務に定常的に用いるデータをまとめたファイルのあるときである。

第 1.11 図 階層構造のあるファイル



この場合はオペレーショナル業務のためのファイルはメッシュ連結をし、上記構造ファイルとオペレーショナル業務のためのファイルの間は複合トリーとなる。

### (7) ファイル・サイズとファイル・シェア

ファイル・サイズを大きくすることは

1. サーチ時間の増加
2. ファイル更新時間の増加(困難さの増加)
3. メモリーの大容量化と処理機器のスケール・アップ
4. ソフトウェア・システムの複雑さ

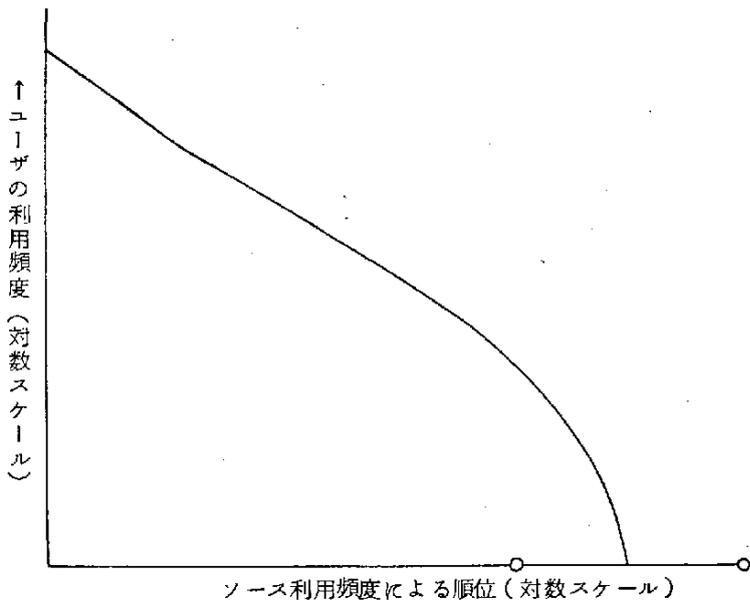
などによって、レコード当りのコスト・パフォーマンスを悪くする。そのため、個々のファイルをコンパクトにすると

1. ファイル・シェアの競合
2. マルチ・データベースの管理システムの複雑さ

が起って来る。

データ・ソースの全く異なるファイル間では、シェア競合の問題は起らないし、同じデータ・ソースからのデータであっても、目的の全く異なるユーザに対するファイル間では、分析・評価・ファイル構成の基準が異なるので、シェア競合が起らない（起ってもトリビアルである。）同じデータ・ソースと同じ利用目的をもつファイル間でのシェア競合の問題は現実問題としては深刻である。

ユーザからみたとき、データ・ソースの利用頻度は対数スケールで直線的に減衰する傾向をもつ。すなわち、データ・ソースを、その利用頻度順に並べると、非常に早く減衰し、その後、同



程度の利用頻度のソースが数多く並ぶという傾向である。これは、ファイル側からみると

1. ユーザの要求をすべてみたすためには利用頻度のよい、比較的小数のレコードと共に、利用頻度の悪い多数のレコードを保持し、維持しなければならない。
2. ユーザの要求によって新しいデータ・ソースを開拓するときは、マージナルな効用を持つデータ・ソースからのデータを処理するため、かえってレコード当りの分析・評価・ファイル構成コストが上昇し、スケール・メリットがない。（スケール・デメリットが増加する。）
3. 同種の要求を持つユーザ・グループをA、B、Cとすると、Aにとって利用頻度の高いデータ・ソースは、Bにとってマージナルであり、Cにとってマージナルであるデータ・

ソースが、Bにとって極めて利用頻度が高いことがある。このときは、ファイルのデータ・ソースの網ら性メリットが大きくなる。

以上のことは、

1. ファイルのレコード構成は、できるだけ利用頻度の高いものにするによって、ファイルの生産性を高める必要性
2. ファイルの網ら性メリットを高める必要性
3. スケール・デメリットを押える必要性

という相互の矛盾する要素の解決を必要とする。これには、

1. 公共的バックアップ・ファイルを、公共財として形成し、ここではファイルの生産性よりも網ら性メリットを重視し、スケール・デメリットを覚悟の上でマスタ・ファイルを構成・維持する。そしてサービス用ファイルにデータを供給する(データはライフ・タイムがあり、時間と共に利用頻度が減少する。この利用頻度が落ちても、公共財としての性格を失なわないものについては、デポジットリー・ファイルもバック・アップ・ファイルの1種として必要である。このときは、ファイルを供給するのではなく、サービス用ファイルから削除したデータを逆に供給してもらって、必要のあるときにとり出すサービスを行なう)
2. 各サービス用ファイルは、バック・アップ・ファイルからの供給をうけた(および、利用頻度の高いものについてののみ自ら収集・分析・ファイル構成を行なった)生産性のよいファイルを維持し、サービスを行なう。

その持つファイルにないデータを要求されたときは、バック・アップ・ファイルに要求を転送するか、それからデータの供給をうけて回答する。

3. 各サービス・ファイルが事故でつぶれたときは、バック・アップ・ファイルから全データを供給する。バック・アップファイルのつぶれる危険性に対しては、各サービス・ファイルからの逆供給をうけて、全ファイルを復元する

という考え方以外に、今の所解決の方法がない。

これは、バック・アップ・ファイルと各サービス用ファイルは完全に重複するし、また各サービス・ファイル間にも、或る程度の重複が生ずる。これは

1. トータル・システムとしての巨視的効率
2. 各サービス用ファイルが、コスト的に成立条件を持つことにより自由競争原理による情報流通産業を形成させる。
3. トータル・システムとしての安全性・信頼性を保証する。

という難点から、必要な重複である。そして公共財としてのバック・アップ・ファイルの形成は

パブリック・セクターの重要な責任の1つである。

[ 理論編 で用いられた技術用語の定義 ]

個体と個体のID： 「何々に関するデータ」という場合の「何々」に当るのが個体である。その個体を識別するために1対1の対応を持つ記号を個体のID (Identification number) という。

属性と属性の値： 個体に関するデータを抽出する場合、その個体を見るいくつかの観点がある。例えば人事ファイルを構成するには、「氏名」「性別」「学歴」「特技」「性格」etcの観点が設定される。これらの観点から個体をみるとき、「個々は、これこれの属性を持つ」という。それぞれの属性に、例えば

氏名	—	○井×夫
性別	—	男
学歴	—	△大学▽学部卒
特技	—	棒高跳び

などの内容を与えるとき、この内容が属性の値である。

アルファニューメリック・データ： そのデータを特性づける属性の値が、何らかの文法や文法的関係をもたない記号で構成されるようなデータを、アルファニューメリックなデータであるという。統計数値、在庫データ、人名録、などである。

センテシアル・データ： 属性の値が日常言語、または何らかの人工言語により、その持つ文法規則によって、「文章形式」をもって表現されているデータをいう。(注1)

アナログ・データ： 属性の値が、絵、図画、写真、立体像のような、アナログ表現によって表わされるようなデータをいう。(注1)

(注1) 論文や記事の場合、本質的にはセンテシアルであっても、それをファイルに収めるとき、文法や文章構成法を無視してキーワードに分解したとき、センテシアル・データをアルファニューメリック化していることになる。同様に、図や写真や分子構造などを、何らかの形でデジタル表現したときには、アナログ・データをアルファニューメリック化したこととなる。

アイテム： 個体に関する情報をファイルに収めるに当たって最少単位となるもので、個体のID、属性の名前などがアイテムとしてとられる。

アイテム・バリュー： 各アイテムの中に収められる、個体のIDの値や属性の値で、特定の記号化法によって記号表現をとったもの。

レコード： いくつかのアイテム（実際にレコードを構成するのはアイテム・バリュー）の集まったもので、ファイル構成の単位となるもの。（コンピュータ用語におけるロジカル・レコードに相当する。）

アイテム・キー： 1つのレコード（すなわち、1まとめになったアイテム群）を識別するのに用いられる記号。例えば個人別にファイルされた人事ファイルでは、個人識別番号か、その個人ファイルの格納番号など。

ファイル： レコードの集まり。レコード内のアイテムの配列法により、シリアル・ファイル（またはダイレクト・ファイル）とインバーテッド・ファイルの2種がある。

シリアル・ファイル（ダイレクト・ファイル）： 1つのレコードに個体のIDまたはそれに代るアイテム・キーを先頭にして、その個体の属性に対応するアイテムが収められているようなファイル。人名録、カタログ集などは、この代表的な例である。

インバーテッド・ファイル： 属性の値またはそれに代る記号をアイテム・キーとし、その属性の値を持つ個体のIDをアイテム・バリューとして構成されるレコードから成るファイル。個人別でなく、学歴別、職階別、職種別等に編成した職員名簿は、この類であり、図書の巻末につけられた索引は、この変型（個体のIDではなく、その掲載頁をアイテム・バリューとしている）

探食用アイテム（サーチャブル・アイテム）： レコードの中で、ファイル・サーチの時に、そのバリューを探索キーとして用いるアイテム。

タグ： サーチャブル・アイテムのバリュー。

1次データ： 収集工程で集められたデータそのもの。図書、論文、新聞の切抜き、記入済の調査票など。これに分析・評価が加えられて、2次データが造られる。1次データが対象とする個体のIDか、またはその1次データの受付番号か、またはその1次データを収める1次データ・ファイルの中でのアドレスが、その1次データのIDとして用いられる。

2次データ： 1次データに対して、その内容を分析し評価を加えた結果を、ファイルを構成する個々のアイテムに対するアイテム・バリューとして表現したもの。そのためには、

- アイテムの種類と個数があらかじめ設定されていること。
- 各アイテム毎に、バリューの集合が設定されていること。
- 各アイテム毎に、バリュー決定の基準が設定されていること。

が必要である。2次データを構成する最少限度必要なレコード内容は、

- (i) 1次データのID
- (ii) サーチャブル・アイテムのバリュー

である。

データ・ベース： データの集まりに対し、それを構成する内容（アイテム、アイテム・バリュ  
ー、その集まりとしてのレコード etc）からみると、「ファイル」という呼び方を  
する。それに対し、利用目的の方からみると、その目的に対して何らかの処理方法（手続）が  
あり、その処理手続きに対してデータを供給するファイルが、必ずしも1つでなく複数個  
あるとき、そのファイルをデータ・ベースという名で呼ぶ。特に個々のファイルを強調す  
るときは、シングル・データ・ベースと呼び、複数個のファイルを強調するときは、マル  
チ・データ・ベースと呼ぶ。

クエリー・データ（Query data）： データ要求があったとき、その内容をファイルから  
求める回答を得るに当ってファイル探索、回答編集が可能ないように記号化、編集したもの。

#### 参 考 文 献

ファイルに処理技術についての解説書として、次のものがある。

中井浩・笹森勝之助等著： 情報探索システム。日本経営出版会刊（1971）

Meadow, C.T.（渡辺茂監訳）： 情報探索—探索言語・情報構成・ファイル処理。

日本経営出版会刊（1970）

Salton, G. : Automatic information organization and  
retrieval. Mc Graw - Hill, (1968)

Lefkovity, D. : File structure for on- syote

( )

個々の実動しているシステムの内容は、それぞれの解説書を参照されたい。

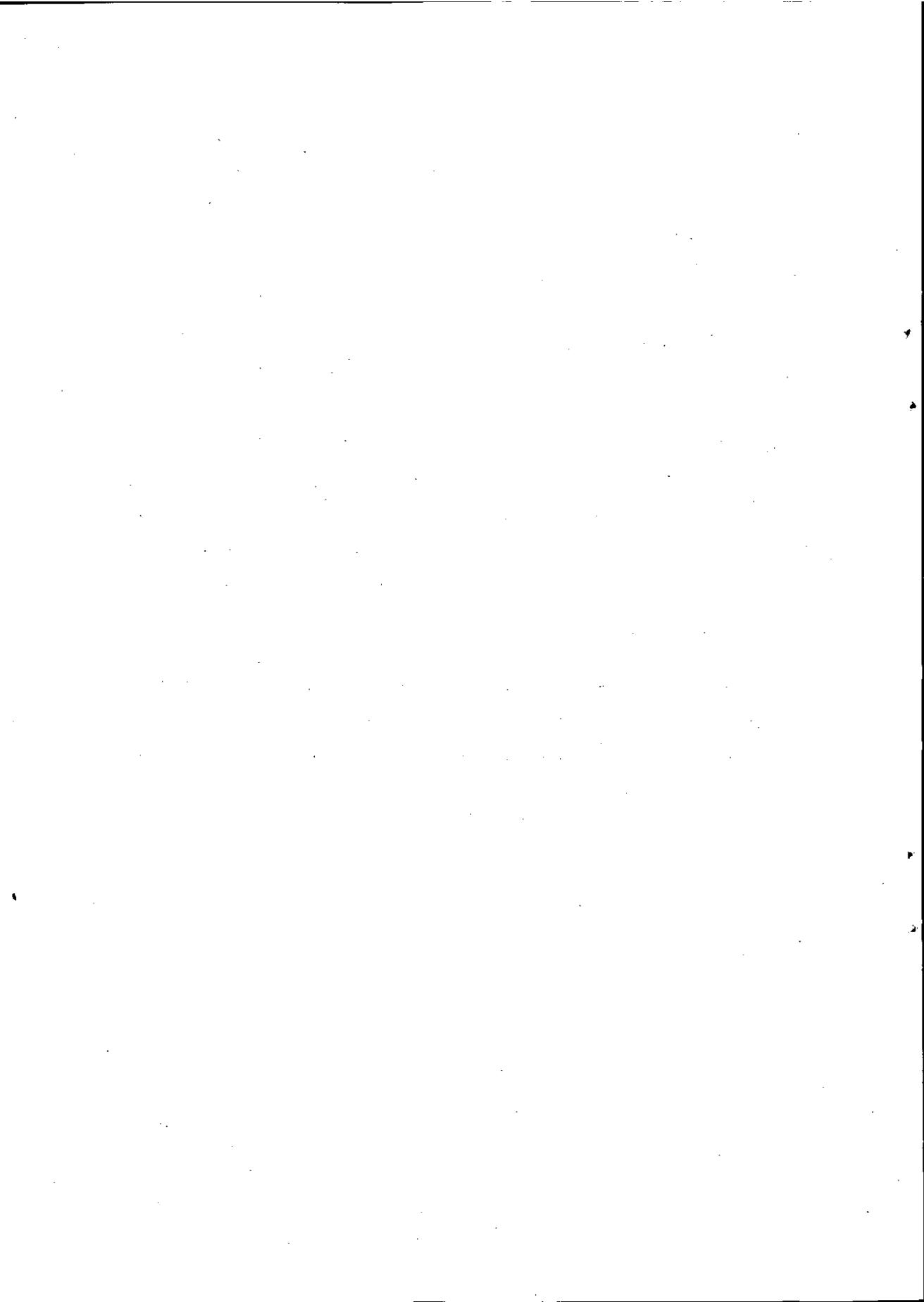
当研究会は、その調査研究の過程において、データバンクの概念規定、その分析の方法論をあらかじめ立ててから事例研究を行なったのではなく、理論的研究、事例研究を併行して行なうことの中から、データバンクという名で呼ばれる事象を如何に把えるべきなのか、データバンクとは如何にあるべきかを、多くの事例の中から模索したのである。この事例研究の成果を、第2部に収めた。

第2部の事例の構成は、先ず最初(2.2)に、当研究会が模索して来た多くの問題意識に対し、極めて現実的な立場から多くのものを、資料としても考え方の上でも提供した、米国加州政府を中心とするデータベースに関するものをおいた。

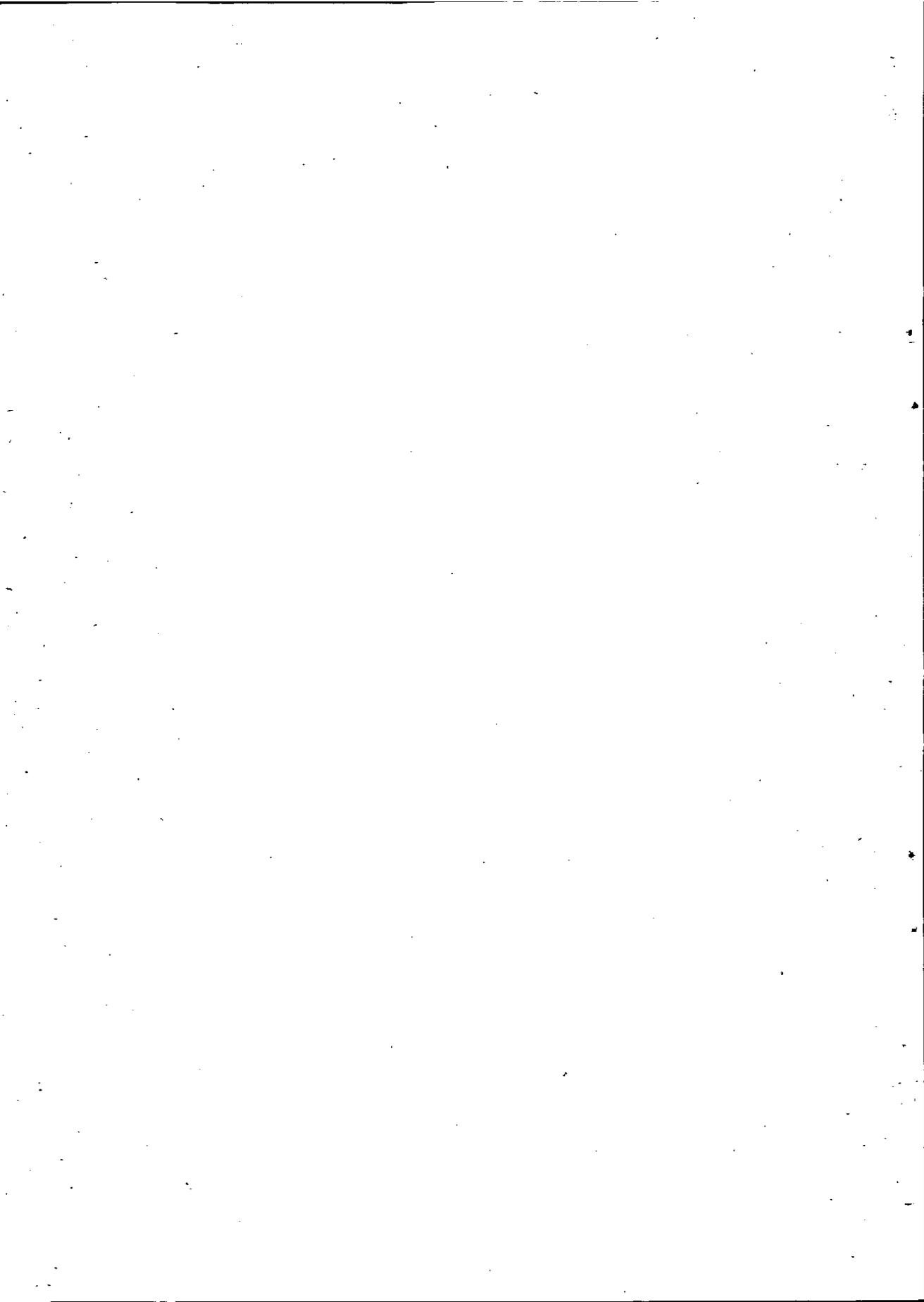
それ以後、米国に派遣した調査団の成果や国内で行なった事情聴取の結果をまとめてある。

これは吾々の事例研究の姿勢が網率的であることを意図しながらも、現実的制約のために断片的、部分的にならざるを得なかった。またこの選択も、必ずしも典型的なもののみを選んだとは言えない。しかし、大体の傾向を把握するのに必要な程度は集めたつもりである。ここに集録されていないものは無数にあり、吾々は氷山の一角を把えているにすぎないという不安は残っているが、第3部の分析においては、その要素を充分意識し、無理な結論を導かない努力をしたつもりである。

この各事例に対する多々の討論結果を、(2.1)に「考察」としてまとめてある。この考察内容が大方の批判に耐え得るものであるか否かが、この報告書のかなめである。それと同時に、これらの事例が、データバンクに関心を持たれる方々にとって参考になれば、当研究会の最大の成果であろう。

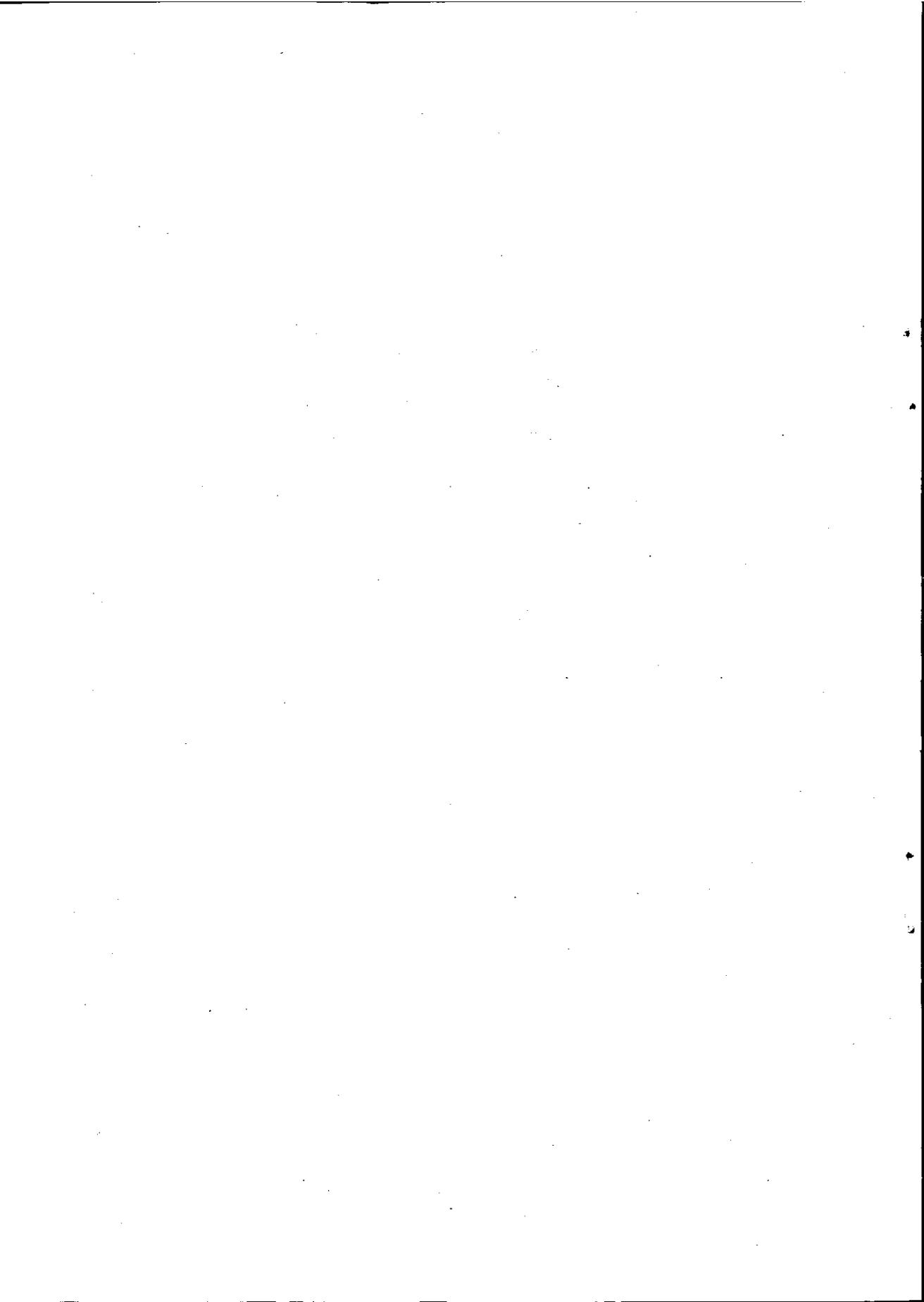


## 第2部 データバンクの事例研究



## I 各事例に対する考察

- (1) 米国加州政府を中心とするデータベース  
に対する考察
- (2) 組織内データベースの事例に対する考察
- (3) 企業性データバンクの事例に対する考察
- (4) 公共的データバンクの事例に対する考察



# I 各事例に対する考察

当研究会は、多くの事例に対する討論の中で、データバンクを模索して来た。その結論とでもいべきものを以下に、箇条書き形式にまとめる。

## (1) 米国加州政府を中心とするデータベースに対する考察

(2.2)におけるマクドナルド氏の概説、討論を通して吾々の受けとったものは、次の通りである。

- (i) 加州政府の中心の箇々のデータベースにおいて基本的な考え方、用いられている技術については、驚くべきことは何もない。むしろ、あまりにも常識的であることに驚いたのである。
- (ii) それは、非常に現実的であり、着実な考え方が基本線として一貫して通っているからである。吾々は、米国連邦政府におけるPPBS導入の過程において、その一環としてデータベースの充実が意識されているのではないかという予測があった。そのため、PPBSと共にトータル・システム意識が先行し、多くの複雑、高度な技術が導入されているかもしれないと思ったのである。しかし現実には、「コンピュータがなくとも当然行なうべき業務」として、多くのデータベースができつつあり、それを根底として、トータル・システムを、遠く見通しつつ、問題意識の根底にしているのであった。
- (iii) その基本線が着実であるのに対し、データベース群全体としてのスケールの大きさ、カバーする範囲の広さ、処理するデータ量の豊富さに圧倒された。人口において日本の $\frac{1}{2}$ 、州政府のコンピュータ保有台数において $\frac{1}{4}$ であるのに対し、コンピュータ利用における費用が3倍というバック・グラウンドから考えるとき、コンピュータ利用の基本的考え方や態度が、体質的に異なっている事を感じる。
- (iv) これは、
  - (a) 州政府の内部業務の徹底的なコンピュータ化。
  - (b) 種々の新しいプロジェクトが、コンピュータをベースとして計画されること。
  - (c) コンピュータといわず、種々の機械や手法を、必要があるとき、現実的な効率判断を基礎として、現実的に導入する米国人の性格。(日本やドイツのように、トータル・システムとか、全体としての効率などを抽象的に考えず、現実として効果があり、現在考えられる範囲が限定されれば、この中で最適システムを考えることは、単に資金が豊富であるからとのみは言い切れない示唆を与えてくれる。などが考え方の背景にある。

(V) データベース網の企画・推進力の中心に、極めて現実的であり情熱をもつスタッフの存在したこと。

このような、現実的な考え方を基盤としない限り、観念の中からは、このようなスケールの大きく、利用面からみて緻密なデータベース網はできないのではないかという印象が、本研究報告を通しての一つの底流を形成している。また、その推進において、中心となる人間の重要性という、極めて人間的な要素とも、吾々は考えさせられることがあった。

## (2) 組織内データベースの事例に対する考察

米国の事例は、調査団の直接訪問した記録を中心とするもので、氷山の一角にすぎない。この訪問先選択において、ノースアメリカン、ロッキードのように、NASAとの関連から、MISの一環として強大なハードウェア、ソフトウェアを駆使してデータベースを形成しているであろう所、ペンシルヴァニア大学のように、自己の研究目的から出発して、他のデータベースとの連携網を形成している所、ボストン小児科病院のように、定まったデータが多く発生し、いくつかの特定目的に用いられる所、そしてニューヨーク連邦準備銀行のように、クラブ的なデータバンクを運営している所など、いくつかの類型によって事例を求めた。これが各類型において典型的であるとはいえないし、また得られた内容にも精粗がある。これに対して、

(1) 一般的なことであるが、一般事務のレベルまでコンピュータを利用する体制ができていくこと。ハードウェア、ソフトウェアを使いこなすレベルが、日本より高いこと。

(2) TSS、オンライン、データ・マネジメントの技術が普及していること。数年前、米  
国を視察した報告において、TSSはまだ普及していないかのような内容を見て行った調査団メンバーとしては、その普及の速度と、「米国は、決してその表面をみてはならず、その根底を支え、急速に顕在化する、その根強い潜在力に注目しなければならない」ことに印象づけられた。

(3) 日常業務の中から、自然に発生して来るデータベースが相互に関連をもち、強力なものに発展して行く過程の中に、「抽象的、観念的な理論よりも実行力」というものを感じる。ということである。箇々の事例について、詳しく知りたい方々にとって物足りない感があるであろうし、類型別にももう少し高いレベルでの調査が何かの機会に行われる必要を感じる。

日本の事例については、MISに関する事例は、現在の所集められなかった。これは、各社の機密に属することであり、当然ではあるが、残念である。その1例として関西電力について、その部分的な紹介を行なっている。財務諸表を中心とするデータの内部利用のためのデータベースとして開銀、長銀の例、株価、財務諸表に対する野村総研の例は、公開して企業的データ

バンクに発展する可能性をもつものの事例である。鉄道の例は、クラブ的データバンクの事例であり、同業種の中でのハードウェアとデータの共同利用を基点として起るものの例であると同時に、この報告の中では表面的には出してないが、データとしては個々の業務を横断的に把握する必要があり、それを利用する例は業務毎の縦割り組織中で利用するというデータの流れを意識的に把握している点で重要である。これに対しては、

(1) 日本は用いている技術では、米国に種々の点で劣っているけれども、これらの事例からみられるように、過去の業績と現実の要請から、着実に成長しつつあるデータ・ベースがあること。この姿勢で、多くの所に、多くのデータベースが育って行くことを強く希望すること。

(2) 内部事務管理において、コンピュータ利用が進んでいる所において、「データバンク」といえるようなデータベースが育っていること。これは、現実的要請と、ファイルを処理する技術がバランスしなければ、健全なデータバンクは育たないことを示している。コンピュータ・メーカの、ファイル処理技術の習熟と、コンピュータの普及と共に、データベースの加速度的普及が予測される。

(3) M I Sとの関連におけるデータベースについては、民間企業の中での意識は極めて高く、計画中の所を含めると殆んどの所で何かの形でデータベースがあるといえよう。

しかし問題は、

(i) 多角経営や、人口の集中による社会の複雑さなど、データ・ソースやファイル収録項目の複雑さにより、データベース・システム設計の困難さ。

(ii) ファイル技術が普及していないこと。

(iii) システム設計のできる人材、それに要するマンパワーが各企業の中で不足していること。

(iv) M I Sは必要だと感じて、データベース・システムに対する現実的要請、理解のないこと。

などが、健全な発達をさまたげている。

### (3) 企業性データバンクの事例に対する考察

米国の事例は、調査団の出発前に、大規模にデータバンクとしてサービスしている所、それ自身としては小規模であっても全米に普及しているサービスの例などを選んだ。この事例研究の重点は、如何なる機構・機能をもっているか、経営的に成立する要因は何か、明らかにすることであった。これに対しては、次のような結論が得られている。

(1) この中に2種の成立条件がある。一つは永い歴史の中で、同種の業務を行なっており、コンピュータの利用により業務の合理化、拡大をして来た所と、もう一つは最初からコンピュータ利用によって出発した所である。長い歴史をもっているものとしてはダン・ブラッドストリート社は120年、ポーク社は100年、クレジット・ビューロー社は70年である。これに対し、スタンダード・プア社は、1967年頃より財務データをサービスするCompustatを商品として売り出したものである。これは最近マグロー・ヒル社に吸収された。マグロー・ヒル社は長い歴史をもつ書店であるが、最近コンピュータ利用による多角的な経営をはじめ、その中にデータバンク業も入っているものである。

(2) 採算性からみると、徹底的な専門化と、徹底的な多角化と2つの傾向がみられる。専門化の例としては、この事例の中にあるものとしては、クレジット・ビューロー社があり、この他に多くの事例の中で、資料にまとめられなかったものの中からとり上げるとデータ・リソース社がある。これは比較的経営規模の小さい所。この要因は、大きいマーケットをもつデータに、努力を集中することによってコスト効果を上げる傾向を示している。この事例の中に上げ得なかった多くの企業性データバンクはこの系列に属する。

これに対し、ポーク社、スタンダード・プア社などは、データ・バンクを主業務とし、多角経営しているものに属する。この種のもは経営規模において大きいものであり、多角経営をすることにより、1つのファイルの多目的利用によってコスト効果を高めることをねらっているものである。これに対し、マグロー・ヒル社のように、主とする書店業務で地盤を造り、それからデータバンクに進出したものもある。このような成立、発展過程は日本においても参考となるものである。

日本の事例としては、組織内データベースの事例としてあげた興銀、長銀、野村総研は、現在、企業的データバンクに移行しつつあるものである。これを企業的データバンクとして眺めると、米国における事例のように長い歴史は持っていないが、業務の中で蓄積されていたデータが、コンピュータの利用と共に、企業的データバンクと発展した例として興味がある。この種のデータバンクとして他にも多くの成立可能なデータ・ファイルをもつ所があると思われる。その意味で、日本における先導的役割を果たすものとして評価すると共に、このような成立要件のあることを強調しておきたい。

紀伊国屋の事例は、まだ本格業務を開始していないがマグロー・ヒル社型の日本における最初の事例として、これもまた先導的役割を果たすものとして、同社の諒承のもとに事例としてあげた。最後に事情聴取の中から、三社ばかりの事例をあげたが、これらはデータバンクといえるか否か問題があるが、日本においては、この段階のものからデータバン

クが育って行く可能性の強いことを考慮して、事例を申しあげたのである。これらは、紀伊国屋も含めて科学技術情報をマシン・リーダブル形式で市販されたことが、成立の大きい条件である。すなわちファイルのみづから構成せず、ファイル内容となるデータの供給をうけてサービスを行なうものである。これは科学技術情報によらず、あらゆるデータについて、豊富なデータ供給があると、民間のデータバンクが容易に成立することを実証するものである。これらを通していえることは

1. 日本は全くの萌芽期であること。
  2. データバンクとして成立する条件は、技術的にも需要の方からみても機は熟しているが、それに踏み切らせるための僅かな加速が必要である。
- (3) その加速は、資金供給であることもあり、技術供給であることもあり、データ供給であることもあるが、もう1つは、「それが必要だ」という需要の顕在化とそれを見通しての決断であろう。

#### (4) 公共的データバンクの事例に対する考察

これは各国に多くの事例がある。特に科学技術情報に関しては無数にあるし、一般情報について、政府機関として多くの機関がある。このため、日本の、代表的な事例として、JETROとJICSTのみをあげて、この中に入っていないが国立国会図書館をも含めて、共通して言えることは、「それを設立すること自身をプロジェクト」とする国家的政策の中で生まれていることである。しかし、事例の中にあげていないものの中に、アジア経済研究所、原子力研究所のように、研究業務の中で集まったデータを、サービスする方向に進む場合も、潜在的には数多くあるものと考えられる。多くの国公立の研究機関は、殆どがその可能性をもつものである。

多くのデータ種類に対し、データを収集しファイルを構成する仕事は、民間では、何かの基盤を持たないかぎりコスト負担が大きいものである。情報産業を育成する上で、資金のみでなく、社会で必要とするデータを、キメ細かく（観念的な抽象論ではなく）現実的に、収集もファイルを構成して、民間にデータを供給する仕事がパブリック・センターの責任であるとすれば、公共的データバンクの設立に、もっと政府は努力すべきである。

また、JETROにせよJICSTにせよ、これからできる特許情報センターにせよ、決して現在の規模では不十分である。（例えばJICSTは、米国の化学部門のみのサービスを行なうChemical Abstracts Serviceと同程度の規模であり、予算的には数分の一にすぎない。）この育成発展に、もっと力をそそぐべきである。

しかし一面、日本の公共的データバンクが官僚化することは避けるべきである。日本の行政

官庁の下にあって、政策の名の下に需要側の真の要求からはずれるサービスを行なうべきではない。それと同時に、民間のデータバンクの育成をさまたげる方向を避けるべきである。民間へのデータ供給、ファイル供給に本質があり、サービスによって利潤をあげる性質のものではない。そのためには資金を、公共投資、先行投資として、これらの公共的データバンクに投下すべきである。

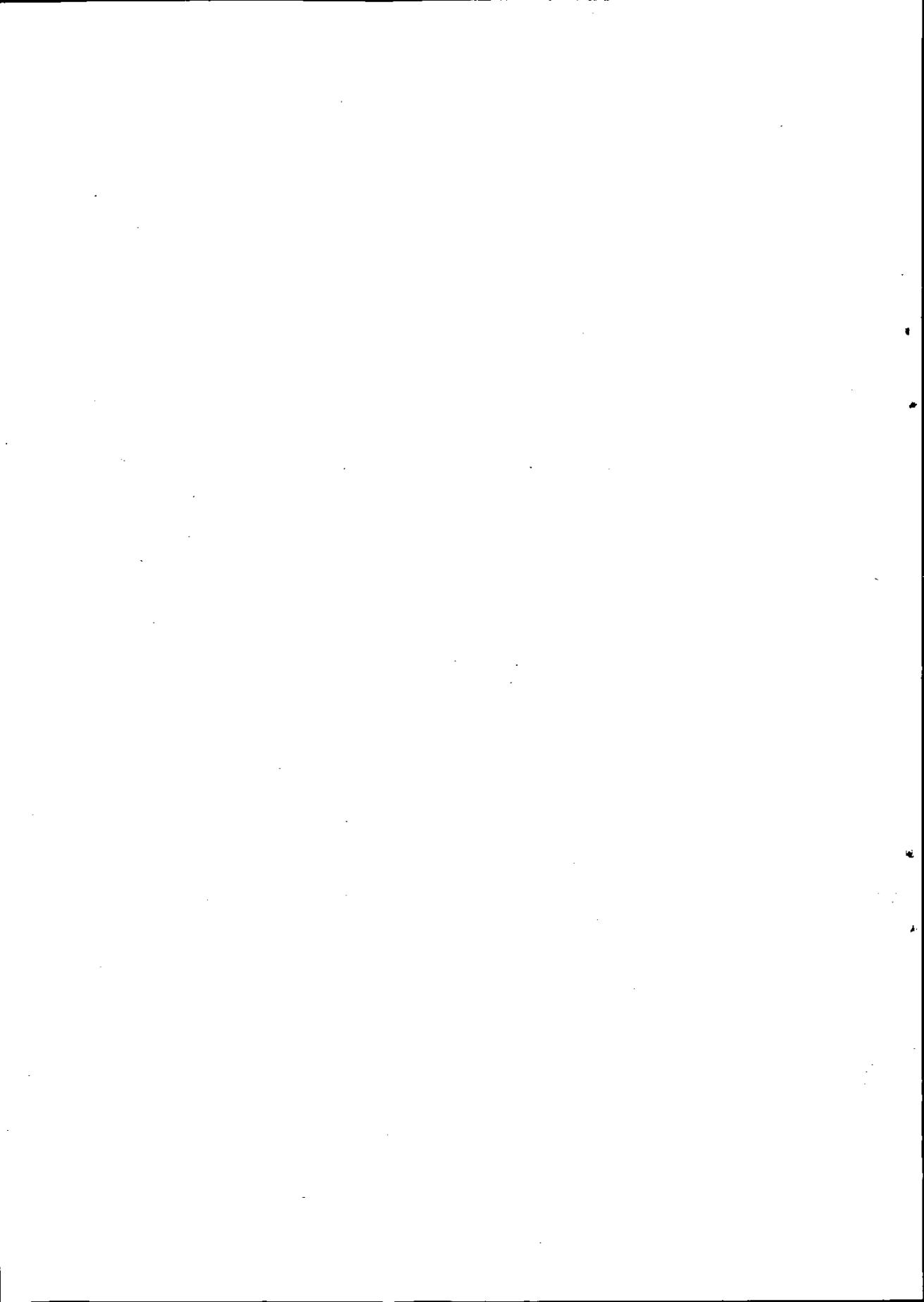
加州で計画した4つの大きいプロジェクト

- 加州集約交通システム（協力会社：ノースアメリカン航空会社）
- 犯罪防止と管理システム（協力会社：スペース・ゼネラル社）
- 廃き物処理システム（協力会社：エアロ・ジェット・ゼネラル社）
- 加州情報システム（協力会社：ロッキード・ミサイル社）

が1965年に出発し、この中でデータベースの構成が行われた。その中から生まれ、FBIおよび周辺他州をも結ぶ所まで発展した犯罪情報通信システムの全容をリストにしたものである。

## Ⅱ カリフォルニア州のパグリック・ セクターにおけるデータベース

- (1) 現状と問題点
- (2) 形成・運営・発展に対する考え方



## Ⅱ カリフォルニア州のパブリックセクターにおけるデータベース

われわれは、1970年3月にアメリカのデータバンクの現状を調査するために、海外視察団を派遣した。その結果、カリフォルニア州においてそれがいちじるしく発展しているのをまのあたりにし、これを体系的に把握することを痛感した。

そこで、この道の権威であるカリフォルニア州財政長官特別補佐官マサイアス・F・マクドナルド氏にカリフォルニア州におけるパブリック・セクターにおけるデータバンクの実態について調査をお願いした。

同氏は、公務多忙のおりにもかかわらず、1971年2月に“Public Sector Data Bank and data processing in California”なる報告書を御送付くださった。

このレポートは以下の目次構成になっている

### Chapter I

State, County and City Governments and  
Public Education in California

### Chapter II

Uses of Data Processing in California

### Chapter III

Observations and Problems

### Chapter IV

Recommendations for Further Study

この内容についての概要は以下に述べる。

#### (1) 現状と問題点

カリフォルニア州は、アメリカ合衆国の各州のうちで最も人口が多く、1970年の国勢調査によると、その人口は約2,000万人であり、その個人所得の総計は800億9000万ドルであるが、州内のいろいろな政府階層（州、郡、市など）の間には人口とか経済活動の程度の違いなどから大きな差異がみられる。

また、いろいろと境界線の違いもある。標準統計区域（Standard Statistical Area）にはそれ自体の境界線があり、州統計区域（State Statistical Area）はまた別の境界線を持っている。更に州には水の問題や水資源供給目的のために地域計画区域

というのが設定されている。

こうした区域に関して多くの問題がある。各区域内の計画を最終的に統合し、州全体としてその他の諸計画と総括して行く上で問題がある。各種計画が互に調和し矛盾なく運営できるようにしたいのだが、カリフォルニア州は未だこうした問題を完全に解決した訳ではない。

市政府間にも大きな開きがあり、カリフォルニア州には400以上の都市があるが、ほとんどのものは非常に小さく、少数のものは逆に非常に大きい。

教育分野では、まづ中央教育局 (Central State Department of Education) が首都サクラメントにあり、州内の教育に関する責任を負っているが、地方の学校の運営については各地方が管理をしている。州内には、1100以上の学校区があり、地方学校区には公選による教育委員会 (School Board) があり、その学校区内に住む人々が選挙によって教育委員会へ代表を送り込み、教育委員会は、その学校区内の教育行政とその意思決定に関して責任を負い、州の教育局は教育に関する幅広い政策を設けることによって、地方学校区を援助するという仕組みになっている。しかし、地方学校区には区内の学校運営に関してかなり広範な自治が許されている。しかしながら、教育補助金のおよそ半分は州政府が出しているが、残り半分が地方政府の財産税で賄われていることから学校区の地方運営委員会は地方財産税の用途について大きな発言権を持っている。

## EDPに関する行政機関連絡委員会

カリフォルニア州は数年前に情報の行政機関間の移転が非常に大規模であり、そして州から地方の政府を網羅するインフォメーション・システムの設計と政策決定の問題に地方政府の人々、つまり、郡、市それに学校区などの人々が発言したいと思っていることに気付いた。

当時われわれは2つの問題を抱えていた。第1の問題は州政府内のいろいろな部局全体の政策方針 (Policy Direction) が不在であり、第2の問題は諸行政機関を結ぶ適当なコミュニケーションがないということであった。

そこで1967年にわれわれは2つの政策委員会を設置した。1つは州データ処理政策委員会であり、これは州の主要な局の長官たち、例えば財務長官、州会計監査人、検事局長、財務庁出納局長、州務長官といった人達に、それに州知事の官房の資源長官 (Secretary of Resources)、人間関係長官 (Secretary of Human Relations)、ビジネスおよび交通長官 (Secretary of Business & Transportation) 農業およびサービス長官 (Secretary of Agriculture & Services) などといった人達から成っている。

その他にもう1つEDPに関する行政機関連絡委員会(Inter-Governmental Board on Electronic Data Processing)を結成した。

この“EDPに関する諸行政機関連絡委員会”(以後IGBと呼ぶ)は州知事官房の4人の長官を代表する4人の州政府職員、郡政府代表の3人、市政府からの3人、教育分野からの3人、それに極く最近になって検事総長の代理として司法執行分野の職員1人を加えたものから成っている。

IGBの設置は大変よいアイデアで、これによって政府各階層間の情報交換が促進した。IGBはそれ自体はインフォメーション・システムを開発する責任を負わず、忠告したり指導したりする責任があるのみで、コントロールしたり認可する権限がないため、委員会が望んだところを全て実現するまでには到っていない。しかし、われわれは実験に協力し協同作業しており、それにプライバシーの保護、コンピュータ情報の安全、標準化の必要性、州の要請を満たす上で地方政府が抱える諸問題等、地方政府レベルの意見に対する州レベルの議会や州知事の注意を喚起した点で大いに業績をあげた。こうした観点にたつてIGBは州政府および州内の地方諸政府でどのようなシステムが運用されており、どのような種類の機器を用いて、どのような種類のデータベースが機能を発揮している、更に将来の計画がどうなっているかを見極めるために上述の調査を行なったのである。

IGBは1か所で成功した経験を他所で生かすことを願う立場にある。つまり俗にいう“車の再発明をする手間をはぶく”ことが目的である。そしてある組織が学んだ教訓を他所へ適用し、知識を他所へ移すことができるなら同じ問題について再び費用を出すことが避けられようという訳であり、IGBは本調査によって各種で何が行なわれているかを調べ、他所にその結果を発表して他の機関での利用を計ったのである。

IGBの技術委員会はインフォメーション・システムのガイドラインのマニュアルを確立した全ての組織でこれを実施せよといった意味での標準ではなく、これはあくまでガイドであるが、もしこれが各行政機関で順守されるならば、やがては互に情報の交換ができるようになるであろうし、情報の様式も統一されたものになるから他の組織が利用することも可能になるであろう。

市や郡や学区などの区域との関連から、またこうした機関の多くが自分達でコンピュータを買い、その設置を別個に始めたために、彼等のシステムは他のものとは独立した個別のものとなり、また独自のシステムを持ったために大きな投下資本を必要とする結果をまねいたのである。

さて、州のレベルから例えば誰れかが来て、“もしデータの記録の仕方の標準化をし、COBOL、プログラミングの標準を設けるなどこうした行政単位間の情報の流れを容易にするために標準化を進めたら、みんなの役に立つのではないだろうか”と持ちかけたとしたら、誰れもかが

“それは素晴らしい考えだ。しかし、私は自分のシステムに既に莫大な投資をし、システムは私が欲することは何んでもやってくれるし、私の目的には結構役立っていますから、今更システムを変える気はありません”と答えるに違いない。これは非常に困難かつ重大な問題である。

私はこの報告書の何処かで述べていると思うが、ある組織はそれ自体非常にダイナミックであり前向きなのであるが、そのことが次の飛躍の妨げになっていることがある。何故なら莫大な投資をした以上、元をとらねばならないので、ある程度の減価消却がすむまではわれわれの提案をのんで大きな変更をシステムに加え、プログラムを修正したり機器を取り換えたりしたからでないからである。

そこで行政機関各階層間の情報の流れがどの程度必要になるかを考えた上で情報交換の標準化をかなり早い段階で行ない、設定された標準に従って地方政府にシステムを作らせればわれわれが抱えているこの重大な問題を回避することができよう。

われわれのIGBはこうした協力と統合を奨励する立場にあるのだが、真の権限なくしてはこうした方向に前進するのはなかなか難しい。

#### カリフォルニア州政府のデータ処理の調査

調査受託会社は主として調査表のアプローチをとり、多くの場合個人面接によって補言した。またこの調査の途中で調査力点の変更が行なわれ、当初調査は開発、運用されているシステムの在庫調べのために開始されたが、途中でIGBはシステムの詳細やデータ・ベースの仕様などより問題分野を明らかにすることにせよという調査方針を打ち出した。

次に日本の県に相当する郡におけるデータ・ベースの用途についての調査結果だが、第1に人口が多い郡はデータ処理予算も大きいことが判る。そして予算の大きい郡は一般的にいって自動化されたデータ・ベースを目指して前進している。

業務の多くは司法執行の分野のものであり、その次が病院管理計画の分野である。病院管理のデータ・ベースの実現に関心が持たれている理由の1つは病院自体がかなり閉鎖的というか限定された分野であり業務の反復率が高いものである。また病院の費用は管理が困難でしかも重要だからである。アメリカでは病院の経費が急激に増加しつつあるので病院管理の向上に大きな関心が払われており、ロックード社を含めて幾つかの会社が病院データ・ベースの運用システムの開発に力を入れて来た。また郡立病院、州立病院それに私立病院でもデータ・ベースの開発に関心がもたれている。

それから、会計や予算などの財務分野でもある程度の作業が行なわれている。それにしても、ここに出てくるようにわずかに4つの運用分野つまり司法執行、計画立案、病院その他の分野しか

開発が進んでおらず、他の多くの政府活動の分野にはあまり着手されていない。

都市の場合にはもっと遅れていて、司法分野に大きな力が入れているが、他は大したことはない。司法や財務分野が最も注目されている。財務分野が何故進んでいるかという点、私の感じでは、財務分野の人々は口頃会計や予算の業務に従事しているので数字に慣れており、こうした方法で情報を得る価値が判ると同時にコンピュータ・システムから情報を得ることに低抗を感じないからだと思う。司法分野に力が置かれている主な理由は2つある。ひとつは連邦政府が司法分野には非常に大きな資金援助を行ない犯罪防止のために出来る限りのシステムの改善を計っていることであり、もうひとつの理由はニーズが明らかなことである。市などの地方地域社会に住む人々が問題を良く認識しておりこの分野の改善を支持していることによる。

### 犯罪情報通信システム

私が知る限りでは連邦から地方レベルまでオン・ラインで動いているのは司法システムだけである。政府関係の業務範囲では司法関係の情報交換機能は断然他を引き離して進んでいる。

### 土地利用情報システム

ついでにもう1つ事例についてふれておこう。数年前にカリフォルニア州政府はいろいろな政府レベルで保管されている土地利用に関する情報についての研究を行なったことがあるが、これはTRWシステムズ社によって行なわれた非常に抱括的な研究であった。非常に広範な調査と質問書による調査が実施されたが、私の記憶ではこれはたしかいろいろな土地利用面に関する基本的情報800種に関するもので、州政府とか郡の評価人事務所、市の記録、民間部門ではローン会社の資料、銀行とか不動産会社など異なる場所で保管されている資料と、その運用システムまで含むものであった。

州の目的は統合情報システムの確立が可能かどうかを調査することであり、このシステムは標準データ・エレメントで構成され土地利用に関する情報、特に土地利用パターンの変化に関する情報を、よりよい計画立案のために政府・民間の各部門間に交換可能にするためのものであった。

われわれは技術的には実現可能であるが政治的に不可能なことを知った。政治的というのは人々には未だこうした企画に協力するまでに心の準備できていなかったということである。司法執行システムの場合には現実には脅威が存在していたので多大な協力が得られ、地方の司法執行機関はすぐさまこれまでの記録方法を棄ててなんの低抗もなく勧告されたシステムに変換できたのだが、土地利用情報の場合は事情が全く違って郡の土地評価人は彼等のシステムについて郡の監督者へも報告しようとせず、郡の財務部長と情報を交換しようとさえもしないという全く

別の世界なのである。

一般的にいて、われわれが抱えている情報交換に関する問題は技術的にはさほど深刻ではないが、変化や情報交換に対する人々の低抗の方が遙かに重大な政治的問題である。

### 自動車情報システム

カリフォルニア自動車局のシステムについて説明するとこのシステムは運転免許と自動車登録に関する完全に自動化されたデータ・ベースを有している。このファイルは単純な運転免許証に関する情報以外のものが含まれている。例えばこのシステムに質問を出して運転免許証システムから次のような運転者に関する情報を要求することができる。すなわち、運転者識別情報、自動車運転上の特典、彼の法的履歴つまり、罰を受けた記録と召喚状を何回無視したとか、事故記録その他の特記事項などであり、これらの情報を組み合わせたものを取り出すことができる。前歴の抄録 (Abstract of Conviction) には違反とその日付、判決とその日付、違反項目、法廷の識別などがのっており、こうした運転者に関する自動車局の記録は一般の人々が考えているものより遙かに完璧なものである。こうした記録は法廷で利用することができ、そこで多くの逮捕歴をもつ乱棒運転、不注意運転、酔ばらい運転などの常習者の記録を判事に提供して判決に資することが考慮されている。法廷には現在のところ端末機が足りないのでオン・ライン・ベースでの利用はされていないが、沢山の端末機をこのシステムにつないで数秒間でこうした情報を検索することが可能である。これはRCAによって設計されたマグネティック・カード使用のマス・ストレージを使用していて、全システムで7,000万キョラクタの情報を収容できる。

第3節の郡政府におけるEDPでは郡における施設の大きさ、CPUのサイズ、データ処理予算と、データ処理を行なっている業務分野は何かとすると、会計とか評価とか会計部門のアプリケーションと司法執行に用いられるものが一番先に手掛けられる分野である。

社会福祉の分野では、2つの郡が社会保障の自動支払システムを開発している。これは社会保障を受ける資格の有無を審査し、ついで自動システムが支払額を決定し、小切手の作成からその送付まで行なり。1つの郡でこのシステムに使っているデータ・ベースの詳細を手に入れることができる。つまり、データ・エレメントとシステム機能の概要を示す。最も遅れている分野は、公共事業とか地方検事事務所とか郡の行政事務所、郡議会などがそれで、この分野でのEDP化が最も遅れている。

技術的要素がEDP・システムの開発を脅かすことがある。技術の新しさからつまり、技術の

進歩があまりにも早いので人々が不安に思い、今実施したなら1年後にはもう時代遅れになるのではないかと考えがちで、いっそのこと今まで通りにやって行って技術進歩が落ち着き、これならという自信が持てるようになったら導入しようとする人達がいるが、これでは何もやれない。困ったことには技術進歩は止まるところを知らないので、適当なところで覚悟を決めてパスに飛乗らねばならない。

もう1つ注意しなければならないことはソフトウェア業者によるソフトウェアの能力についての誇張である。これはいくら強調しても強調しすぎることはない。

私個人の経験でもあるしあらゆる調査関係者が経験しているところであるが、コンピュータ・メーカーやソフトウェア会社の人達が彼等のソフトウェア・プログラムやパッケージの能力について云うことをうのみしないよう、よほど用心しなければならない。

殊に契約の際に慎重を期さねばならない。契約条項、条件、罰則などをきちんと規定して、彼等の計画について法律的责任を課すようにしなければならないのである。

例えば、予算データ・システムの開発の際に、(これは私がプログラム・ディレクターをしているシステム開発であるが)われわれはデータ管理のための特殊なソフトウェア・プログラムのデータ・マネジメント能力を使用しようと計画した。われわれはダイナミックなデータ管理能力を備えて、何時でも自由ファイルを追加し、記録やファイル・オーガニゼーションも自由に変えられるようなものにしたいと考えた。つまり完全なデータ・マネジメント能力を得ようとしたのである。

私が知っている限りで、そういうようなプログラムがあった。それはDATA LANGUAGE Iと呼ばれるものであった。予算システムの開発のためにわれわれは非常に有能なコンサルティング会社と契約を結んだが、彼等はDLIを用いることによって希望通りのデータ管理能力が得られると推薦した。彼等の忠告を信頼していたのだが、実際に使用する段になってDLIの1種で権利無効状態(Public Domain)になっているものがあることが判った。これはIBMから入手できるが、IBMによってサポートされておらず、もし事故が起ったらわれわれが自分達の手でどこが悪いか診断しなければならぬというものだった。

もう1つ別のDLIの変種があった。これはノースアメリカン社のもので自家用に改良し使用しているものであった。このことを発見してノースアメリカン社と交渉したが、売ってくれるだけでなく人をよこして設置を手伝ってくれるとのことだが、値段が10万ドルというので折合わず、結局われわれは中間的処置としてより小さなデータ・マネジメント・システムで我慢しなくてはならなかったが、この間6か月という期間を無駄にってしまった。

要は専門家であるコンサルティング会社の勧告に基づいてこういうことが出来る筈だというこ

とで始めたのに、実行段階になって駄目になってしまったということである。

ソフトウェアの問題に関するもう1つの例は、カリフォルニア司法執行通信システム(CLETS)の開発の際に、基本的な機器の契約を得たRCAがこれに用いられるソフトウェアの開発に当った。RCAはサブコントラクターとしてINFORMATIC社を採用しソフトウェア会社の同社にプログラミングを依頼した。インフォマテック社は有能な人材を抱えた優秀なソフトウェア会社なのだが、必要なプログラム開発設計に対する十分かつ効果的なマネジメント・コントロールに欠けるところがあったので、1年計画実施とか運用段階を終了する筈であったのに、プログラミングに掛ってから9か月位してから、とても予定期日までには仕上りそうにないことが決定的となった。9か月の時点でインフォマテックは予定より6か月以上も遅れていたのである。

さてこの計画の場合、われわれは期限に間に合わなかった時の損失に対する補償項目を契約に入れてあったが、この補償項目(Clause for damage)は大した金額ではなかったが契約にのっていたので遅れた場合には支払を停止するばかりか逆に賠償金を支払わねばならないことになっていた。結果としてわれわれは損害を補填し、スケジュールをやり直し、実施を6か月延期せねばならなかったし、その上RCAとインフォマテック社は期日に間に合わせるためにおよそ100万ドルの余分な金を支払わねばならなかった。この作業を遂行するために何が必要かについての見積りが甘かったためである。

要は自分でシステム設計をしプログラミングなどをするならばともかく、外部の会社にやらせる場合には別に人の悪口をいうつもりはないが、しっかりした契約を作成しておかなければならぬということだ。何故なら、この分野では人々は善意の間違いを犯すことがあるが、こうした間違いに対して自分自身を守らねばならないからだ。

建設業などの場合には賠償責任の項目を設けるのが慣習になっているが、コンピュータ分野で弁償責任の契約条項を設けたのはカリフォルニア州政府が初めてである。

時々会社がわれわれのところへやって来て、あなたはこれができるとかあれができる、こうしたらどうか私達を庸ってくればやってあげるがなどと勧誘に来るが、本当は彼等は“これは既に行なわれている。私は未だやったことはないが、私にやらせてみてくれないか”というべきである。どうもコンピュータ分野では誰れかがやったことは容易にできると思ひこむような風習があるが、実際にやってみるとそうは行かないことが多いようだ。

もちろん経験のある人の数は限りがあるし、データ転換や入力の速度には限界がある。例えば自動車登録とか運転免許証のシステムなどの場合は、手書きの記録を電子制御様式に転換するのにカリフォルニア州は4年の歳月を要した。殊に2つのシステムを運用して行く、つまりあるも

のはコンピュータ・システムの中に既に記録が入っているし、あるものはファイルの引出しに未だ残っているといった具合に2つのシステムを平行して行くのは非常に困難なことであった。データ転換は綿密な計画とスケジュールとをたてた上で慎重に行なわねばならない。

カリフォルニア州にある多くの都市は人口も少なく、大規模なEDPの利用も行っていない。自分達のコンピュータを持っていない都市は計算センター( Service Bureau )を利用している。

MDS ( Municipal Data System ) と PADS ( Public Agency Data System ) は2つの協力事業であり、大変興味ある実験だと思ふ。いくつかの小さな都市が集まって合併事業をしようというもので、これは共同のコンピュータ・システムとデータ・ベースを作ることを目的としている。もちろん共同のデータ・ベースといっても個々の市町村の記録は別個に記録されるのであって、こうした努力によって個々にはコンピュータシステムを持ってないような小さな自治体がシステムを持てるようになるという訳で、これから2つの実験の進展は注目してほしい。MDSの方がその計画や実施においても1つの実験よりも遙かに進んでいる。

### 学校区の情報システム

学校区で利用されているシステムについて述べよう。カリフォルニアの学校区ではいくつかの努力が行なわれている。その面で報告様式の統一が要求されるようになってきている。州の教育局には州としての学校区運営計画を各学校区に割り当てる責任があるが、あまり報告書の様式が複雑なため — 実際この様式が判っている人は州政府内でも2~3人しかいない位だ(笑) — これらの様式を廃止して新しい簡単なものを作る必要があると考えている。何故そのように複雑かという州議会が毎年のように法案を提出し修正するので修正の上に修正が重なって読んでも判らない位だ。州教育局は予算を握っているので相当の力を持っているので、州教育局は地方からの報告書の報告事項の変更を行なり。そして同局はCEIMS ( Califonia Education Management System ) データ・ベースと呼ばれる記録の蓄積をこの数年に渡って行なってきた。この種のデータ・ベースの要素とどのような種類の情報が入っているかを述べた。

私は個人的にこのデータ・ベースより、情報を用いて仕事をしている。昨年から州財務局は州内の公共教育システムの研究を行なっている。これは学校教育の現状を把握・評価し、そしてどのような要素が実質的な機能の違いを教育分野にもたらすかを見出そうとしているのである。つまり社会経済的特性すなわち、学生が教育課程に何をもらすか、彼が貧困家庭の出身か？ 彼がアメリカで困難に直面している少数民族社会、例えばメキシコ系の子弟か？ (時にはアジア系の人々も困難に直面することもある)、彼の家庭はどうか、離婚家庭の子弟か？ 住居はどうか？

文化的背景はどうか？ などの情報を得た上で、彼の行っている特定の学校へ提供しうる資源の状況を調べる。(これは学校区によって貧富の差があるから)それから各学校で実施されている標準学力テストの成績もある。その他実際に学校区がどの位の予算を何に用いているかの記録もある。

われわれがやろうとしていることはこうした諸要素の関連分析を行ない、何が高い機能を発揮するのか、そしてそれらの関連はどのようなものかを見出すことである。もう数年にわたってこのようなデータ・ベースの作成を行なっているが、完成までにはもう数年を要するであろう。

### データ・バンクに対する所見

カリフォルニア州におけるこうした活動の現状について所見をのべてみた。そしてデータ・バンクとデータ・ベースの違いや統合情報システムとの違いについて述べた。

最高の技術知識の投入と創造的想像力、管理能力(management skill)、巨額の資金が効果的なデータ・バンクの運用には必要である。中途半端なことをやると失敗し、人々を失望させるぐらいなら着手しない方がましだと思います。こうした分野は1つ1つ重要なものであり、何が可能であるかについての技術的知識、新しいことを試みたり同じことを異なる方法で試みたりする創造的想像力、全計画のスケジュールをたて管理し、期限と予算内に納め、困難な意思決定を行なり管理能力がなくてはならない。

困難な意思決定とは、あらゆるシステム開発はある時点において、次から次へと出て来るアイデアを押えて、一定の構想の下に期日内の完全をめざして作業を進めるという決定をしなければならなくなる。これを「設計の凍結」と呼んでいるが、これはマネージャーが時間と予算を考慮に入れて断固として行なわねばならない。

また、経費のことであるが、大規模なプロジェクトの場合には多額の金が無駄になる。自分としては「無駄」になるという表現はしたくない。意図したことに費やされると云うべきであろう。しかし私の16年間の経験からいって、何十億ドルという金が研究に費やされる。つまり何十億という金が「これをしては駄目だ」ということのために使われるのである。人々はとかくどの位の金額が「これをしてはならない」ということを発見するために費やされたかを忘れがちだが、実際にそうなのである。そのためにこうした大きなシステム開発には巨額の予算が必要なのである。しかしその経過を知らない人達は何故最初からそうやらなかったのだと責めることがある。

## (2) 形成、運営、発展に対する考え方

前節において、カリフォルニア州のデータバンクの現状に焦点をあて、その成立の背景や評価をおこなってきたわけであるが、この節では、データバンクのあるべき姿とは何か、また、それをいかにしてつくっていくかなどを考える上での事例として、カリフォルニア州の考え方を追ってみよう。

われわれは、情報公社とか情報省というものがあって、そこで一元的に情報をとりあつかうのはデータバンク全体としてあるべき姿ではないと考えている。それは実現不可能であるからである。むしろ相互にメリットなり特性をもったデータバンクが沢山でき、そして情報の内容が相互にからめあって、多角的に利用できるいわゆるネットワーク構造こそデータバンクのあり方として望ましいし、またそのようにするのもっと現実的かつ本質的であろう。

この件に関するマクドナルド氏とわれわれの討論の記録が、次頁以下に述べられている。

この討論の記録をみれば、マクドナルド氏の豊かな経験に照らして述べられた意見の中にデータバンク構造はトリー構造のそれなのか、ネットワーク構造のそれなのかもまたおのずから明らかとなる。

### 〔質問1〕

マクドナルド氏は、データバンクをどのように考えているか？

### 〔回答〕

データバンクは、ただ単に機能的ファイルの情報によって構成されているのではなく、それ以外にも将来なされるべき計画や決定に関して、データバンクの概念を適用するような目的も持っている。従って、市場調査の結果出来たような、機能的なデータ・ベースにもつながりを持ち、さらに基本的な経済的プロジェクトとか、企業の運営にGNPが将来どういう影響を与えるかなどの情報も含む。データバンクの概念は全体的オペレーションを頭にいれて綿密に設計されたものであるから、いついかなる時、どの機能的データ・ベースの中にある情報が必要になろうとも、広範囲な主要目的のために使用されることができ、しかも、それらデータ・ベースにある情報は、何らかの修正を加える必要なく使用することができるように設計されている。

何年前かに私の国（アメリカ）でしばしば使用された言葉に“トータルシステムズ・アプローチ”というのがあるが、このデータバンクの概念はこの言葉の精神を適用したものと、ある意味でいえると思う。つまり、ある与えられた機能的データ・ベースの範囲内で何をしようとも、その情報源からの情報をうまく組合せて、より広い目的のために使用するというのがこのデータバンクの意味である。

このトータルシステムズ・アプローチの考えには、非常に困難な又、非常に現実的な問題があり、よく起ることだが、例えばただ単にトータルシステムズ・アプローチの観点から問題をながめると、そこから出て来る提案というものは、あまりにも広範囲であり、統括的であり、非常に経費がかかり、実現させるには時間がかかりすぎるため、そのようなシステムに対して疑惑の念を抱かざるを得なくなるし、トータルシステムズ・アプローチが基となったこのような莫大なスケールのプログラムを一般に売り込むのは無理だと思わざるを得なくなる。従って人々がそのようなほう大な仕事をする能力を疑うばかりでなく、トータルシステムズ・アプローチ自体に含まれる個々の機能にもすぐに解決しなければならない問題が多々ある。しかし、人々はトータルシステムズ・アプローチが完成するまで待っていることは出来ない。従って、人々はシステムズ・アプローチが完成し、自分達の機能をシステム化するのにそれを適用することができるまでの間、現在の機能的な必要性を満たしてくれて、他のいろいろな問題を広範囲に扱ってくれるようなシステムを望んでいるのである。

他方、システム作業が、全体的な設計概念なしに個々の機能部門で別々に進められるならば、それぞれの個々のシステムを後に統合することは、非常に困難になりがちである。というのは、個々のシステムが他の関連部門の必要性を考慮に入れずに、特定の目的のために動くように作られているからである。工場を例にとると、工場長は会計システムを考慮せず自分の仕事のコントロールシステムを作ってしまうがちであるが、もし労働賃金や部品価格を会計システムと切り離れたらば、その企業はコントロール不可能となるであろう。

これらは困難な問題で、そうやすやすと解決できるものではない。我々が開発中の概念を、カリフォルニア州政府からいくつかひろって説明する。

つまり、カリフォルニア州政府の運営を全体的に見る我々と、ある一部門のみに関連する人々と対応させて説明する。

まず考察のために矯正部 (Dept. of Corection) をとりあげる。これはかなり大きな部であり、カリフォルニア全体には多数の刑務所があり、その囚人数もけたはずれのものである。機構がほう大なだけにその機能もさまざまで、インベントリーを持ち、購入をし、維持活動等も行わなければならない。又、囚人に関してもなすべき仕事は多い。まず、第一に法廷で判決が下されると、新しい囚人を受け入れ、囚人に関する記録を作成し、どの刑務所に入れるかを決定し、囚人を分析して、中学卒程度の者には高等教育を授けたり、職業訓練が必要かどうか決定したり、いろいろな心理分析をしたりしなければならない。つまり囚人を診断し、その結果、社会復帰を助けるわけである。さらに囚人に関してのプログラムに対しいろいろの決定がなされる。つまり例えば、高等教育を受ける囚人数が昔に比べて急に多くなったとすると、高等教育プログラム

(もちろん刑務所内での)を拡張しなければならない。従ってそのようなプログラムの監査と評価の決定がなされるわけである。そして、囚人が仮釈放の資格を得ると、全く別のプログラムが用意されている。囚人は刑務所からは出られるが、仮釈放プログラムの下に監視されていて、もしこのプログラムの基準を満たすならば、最終的に釈放されるわけである。

さて、この部はこのようにいろいろの機能をもっているが、それらは必ず何らかの記録保持と関連している。がこれらは、全く別々のデータ・ベースである。そして我々は今この部をトータルシステムの実行可能性を決定するために研究しているが、もし設立されたならばデータバンクとして働くようになるだろう。しかし、個々の部で起っているすべてのことは他の部と関係していることからして、これは、単にデータ・ベースでしかないとも言える。

さて、今ある男が法律を破って、店又はガソリンスタンドに押入り、窃盗をはたらいて、車を盗み、逃走したと仮定する。そこで、施法官が探査を始める。逃走した男は、前科があるかもしれないが、もしあれば、FBIにそのファイルがあるであろう。そして我々はその男に関するファイルを持っていないとする。又、ハイウェイパトロールは盗難車の連絡を受け、その車の特徴や、車輛番号などの報告を受ける。もちろんこれらの情報は、G.L.E.T.Sシステムを通して得ることができる。さてこの男が逮捕されてその記録が作られ、彼に関する情報が集められ、最終的に起訴され、その裁判に関する記録が作られ、判決が下って、この男が刑務所に入れられると仮定する。

この場合、この犯罪ケースについての情報のすべてに関して、一つのデータ・ベースの形で考えるのはあまり実際的とは言えない。このケースに関するすべての情報が決定を下すのに役立つような形でシステムのどこからでも得られるような形であるべきと思われる。

もし、この男が刑務所に入ったなら、彼の家族は社会福祉を受けるだろうし、又その資格があれば医療サービスも受けられ、子供の一人が麻薬中毒で、精神衛生部のお世話になるかもしれない。又彼の家族の他のメンバーについても、職業斡旋がされるかもしれない。このように彼に関しての決定は矯正部が行なうが、例えば、もし彼が仮釈放の資格があるとみなされると、彼の家族に非常な影響を与えるだろう。もしこの男が釈放され職を得て家族を養うことになる、家族は今まで受けてきたいろいろな福祉プログラムを受けなくなるだろう。つまり、その釈放の決定が非常に重要な意味を他の分野にもたらすわけである。他方、もし釈放されないならこれらの福祉は引続き行なわれる。

現在、我々はこの男に関する知識と他の別の分野の記録とを結びつける方法を持っていないので、すべての段階における意思決定は不十分な知識に基づいて行なわれている。再度強調するがそれらは大変大きな機能分野であり、医療補助プログラムはカリフォルニアだけで一年に10億

ドル以上かかるし、福祉プログラムは一年に5億ドルもかかる。これらは非常に綿密に作られた管理プログラムであるため、一つのデータバンクのみでこの男もその家族に関するすべての情報を扱うと考えるのは不適當であろう。しかしこの男に対して決定を下すときに必要などんな情報でも得られるようなコミュニケーションリンクや情報のコード化というものが確立されねばならない。従って現在我々はすべての分野に共通するデータバンクよりも個々の分野のためのデータバンクについて考えているのだが、その目標として、“統合された情報システム”(integrated information system)をとりあげている。これは、各々のシステムにある情報関係を見出し、これらの間のコミュニケーションリンクを作り、決定時にすべての必要な情報を得られるようにするものである。又これは、情報がやたら乱用されないようにコントロールされねばならない。大変困難な問題であるが、我々は今日その解決に向けて努力している。

〔質問2〕

マクドナルド氏は、公共部門でつくられたデータ・ベースが民間で利用するという型に広がる可能性があると考えるか？

〔回答〕

現在までアメリカでパブリックセクターとプライベートセクター間でデータの分割と交換をするための協力体制というものは特定の意味においては全くといっていい程ない。しばしば共同情報システムの設立案が出されてきたが、それは例えば、銀行、金融業、不動産会社が政府と共同し、経済予測等に使用する特定情報を蓄積するものである。私の知っている限りでは、カリフォルニアにおけるこのような提案は今まで成功したためしがない。しかし、違う意味での政府と私企業の協力や、政府によって定められた情報を私企業がアクセスすることはたやすいのである。

我々は、伝統的に政府は人民のものであるという信条をもっている。言いかえれば、どんな人民も政府の仕事を理解し、その内容について聞き、行なわれたことすべてを知る権利を有する。又人民は納税者として払った税金によって行なわれた政府の仕事の結果から得られる利益をもつ権利をも有する。

このような態度は広範囲にわたり、私企業までも個人と考えるようになった。つまり、もし公的な部局が広範囲な記録を集めたとする。例えば車の持主名と住所、免許所有者名と住所、不動産持主名と住所等に対してどんな人でもそれが政府によって集められ、まとめられたものであっても、伝統的にこれら記録情報を得る権利をもつ。ただし個人のプライバシーを侵害しないという条件を守ればのはなしである。そうすればこれらの記録はただで、コピー代を払うだけで得ることができる。

この原則は個人の情報を犯すところまでは適用するわけにはいかない。例えば犯罪歴、精神病院での記録のような情報までには及んではいけない。しかし、その他の情報は個人にでも企業にでも公に使用できる。従ってある意味で我々は政府から一般のプライベートセクターへ大変オープンな情報の流れをもっているといえる。しかし初めに申し上げた通りに他の意味でパブリックとプライベートの両セクター間での情報交換の協力はあまりない。これはそれほど簡単にすまされる問題ではない。公的情報を私企業に営利目的に使用のため与えることに非常に強い反対が起っている。それに対し我々はいつも誰にも、どの企業にも不公平なく情報を与えるつもりだと解答しているが、それで皆が満足するわけではなく、なぜならその情報を利用する資源に差があるからである。更に個人の名前と住所のリストが社会的にあまり好ましくない商品を郵便で勧誘するのに用いられたことがあった。しかも情報要求の内容の良し悪しを選り分けるのは非常に困難である。従ってこの原則を実施するのに問題はつきものであるが、しかし我々は今だに政府の記録を私企業に自由にアクセスさせるようにしている。

〔質問3〕

多種類のデータ・ベースを組合わせて使用するために、それぞれのデータキャブラリーをお互いに一致させていると思うが、警察の取締りや、事故記録のシステムで使用しているパーソナルコードが、裁判とか免許にも通用しているか、それは又自動車事故に関する民間の保険会社にも共通になっているか？

〔回答〕

我々は、汎用個体識別コードは持っていない。実際にデータ処理等に関して識別のために汎用されているのは、住所、氏名、年令、性別のような基本的情報である。これらを用いて個人の記録を得る手がかりとするのである。多くのデータ・ベースにおいて識別コードやコード・ヘッダーを用いて情報をストアしている。しかしそのコードは、そのシステム内でのみジェネレートされたものであり、氏名の文字、住所の文字数、性別や年令要素を用いて、ある特定のソフトウェアプログラムにより作られたものである。しかしそのような識別コードは、そのシステム独自のものである。従って基本的には、氏名、住所、年令などを用いて識別しなければならないのが現状であり、統一された汎用識別コードはない。統一識別コードに最も近いものは、社会保障システムの結果であるといえる。社会保障は連邦政府によって行なわれている老人、不具者のための保障であり、ある人が初めて雇用された時社会保障口座番号（social security account number）が与えられこれは一生用いられる。連邦政府に関しては、その人についてはその社会保障番号が用いられる。ワシントンの社会行政部で、彼のための口座が設けられる。この

口座に、彼の給料や雇用人からのお金が積立てられる。これによって彼が停年や不具になった後、生活にことかかないようになっている。決してその番号は変えることはない。連邦政府に報告される所得税は、この社会保障口座番号を用いている。従ってこの番号は標準的の番号といえる。たびたび、はっきりと標準番号にするという提案がなされてきて、例えば子供が生まれたら、すぐに社会保障番号を与え、一生涯その人の記録、学校成績、銀行口座までも、それによってコントロールすべきであるというものである。もしこのような統一コード番号がユニークに決まれば、数々の利点があるだろう。

しかしこの考えはまだ一般的に受け入れられていない。将来受け入れられるかどうか疑問である。統一コード番号は数々の欠点もあり御承知のように多くの企業では、それぞれ独自の識別番号を作っているため、これを新しい番号に変えるのは、あまり気がすすまないはずである。更に社会保障番号は完全に認められた (fully proof) 番号ではなく、もしある人が自分の身分を変えたいなら、違う名前を使うことができるし、そうすると社会保障番号がないからといって、新しいのを要求することすらできる。しかし利点もあることもある。まだ統一された唯一の認識番号というものができかどうか分からない。

#### 〔質問4〕

個別データ・ベースと中央大型データ・ベースとの相互関係をどう考えるか

#### 〔回答〕

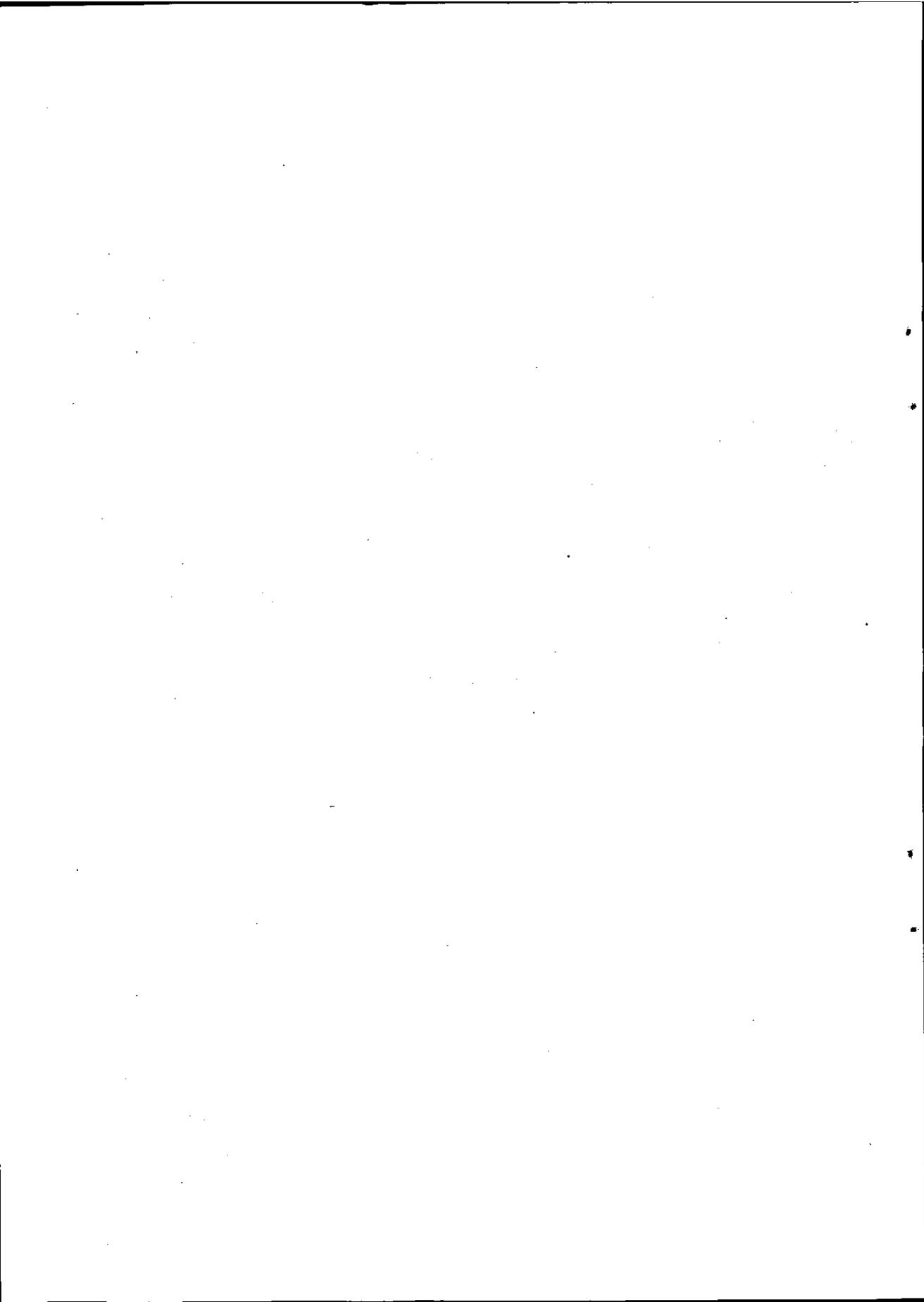
これに対する答えは、容易には出せないが、規模の経済の尺度を適用することは不可能と考えられる。つまり、小さな個別の組織より強力な中央組織によって、記録を更新する方がコストがやすいわけで、この経済要素は一つの基準となる。しかし、独自のデータ・ベースを(データバンク)もっていて、他のすべての部にサービスをするのは、あまり実際的ではない。また、あまりにもシステムが大きくなり、それを使う人からかけはなれたものになり、仕事に対して関心をなくしたり、情報が更新されているかどうかも見ず、仕事の役割りを果たそうとしなくなる。従ってシステムはそれを使って働き、そのサービスを受ける人達の身近になければならず、あまりかけ離れると、お互いの機能がうまくいかず分離してしまう。このシステムと人々との距離がどの程度が良いかはわからないが、使用者がブレイクダウンしないように一諸に働ける位が良いかと思える。

〔質問5〕

カリフォルニアのデータバンク開発の基本は何か。

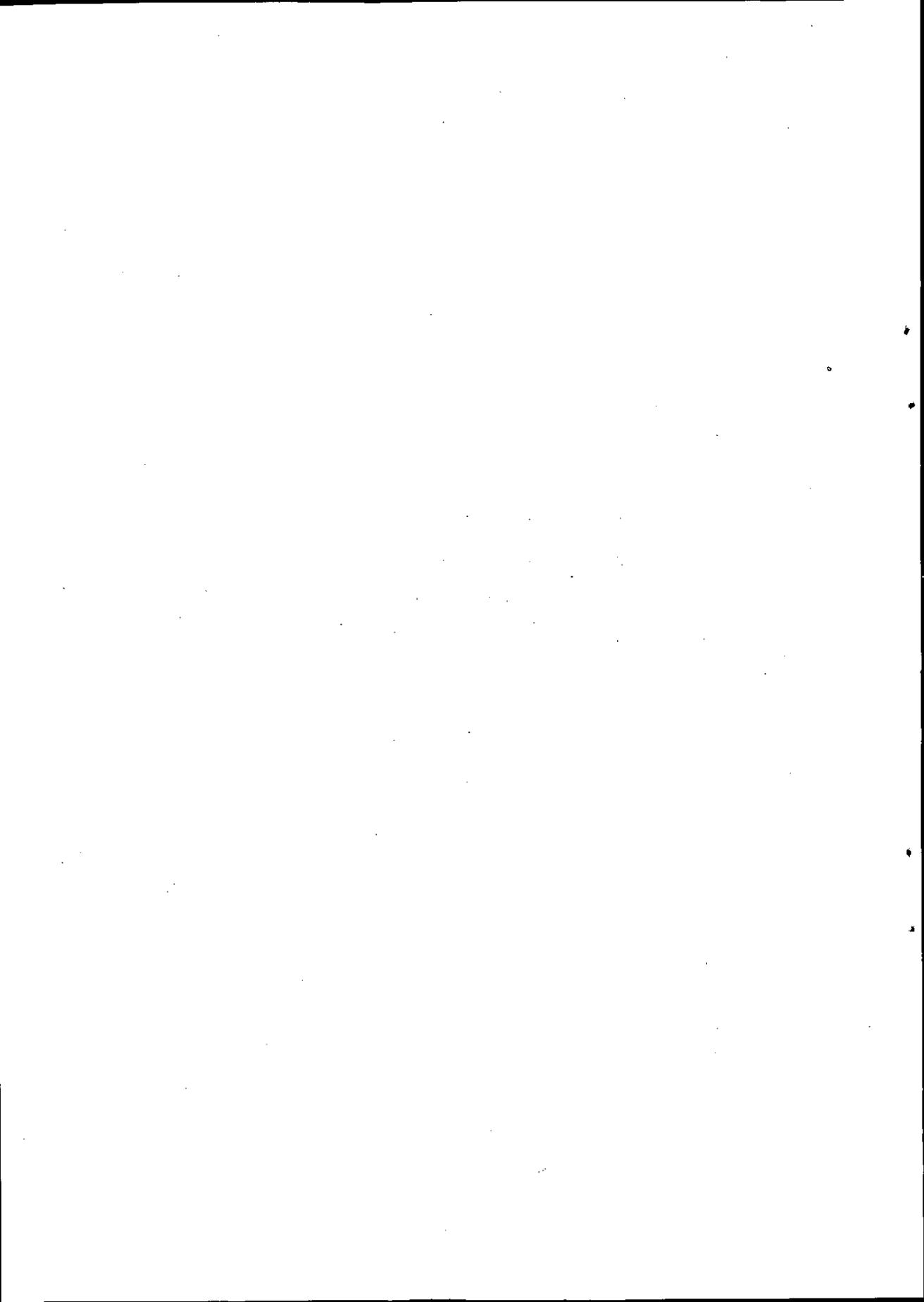
〔回答〕

このような大型開発プログラムの基礎は一般的には、あまり効率よくなかったが、すでにそのような機能が実行されていたということである。統合システムを開発する提案はいつでも現在のシステムを改良する提案に他ならなかった。改良には費用がかかるが、しかし長期的にみると経済的である。というのは、現在の手計算よりも安いユニット当りのコストで同じ仕事ができるからである。また、問題が大規模で重要性を帯びてきたとき、コンピュータのスピードを用いなければどうしても解決できないというようなケースもよくある。例えば、カリフォルニアでは2000万台もの自動車の車輛登録、1400万の自動車免許記録のようなほり大な量を維持更新していくには、どうしても人間の仕事ではすまされないとこまできているのである。州政府の例をあげると、州が管理している教職員退職システムの記録があるが、これらの記録は二年も遅れており、これを自動化して、退職時に口座預金がどの程度あるかの質問に対し、二年前の記録を用いないよう、現在の残高を答えられるよう努力している。



## Ⅲ 組織内データベース

- (A) ノースアメリカン・ロックウェル社
- (B) ロッキードミサイル&スペースカンパニー
- (C) ニューヨーク連邦準備銀行
- (D) ボストン小児科病院
- (E) ペンシルベニア大学
- (F) 日本開発銀行
- (G) 日本長期信用銀行
- (H) 関西電力
- (I) 野村総合研究所
- (J) 鉄鋼連盟

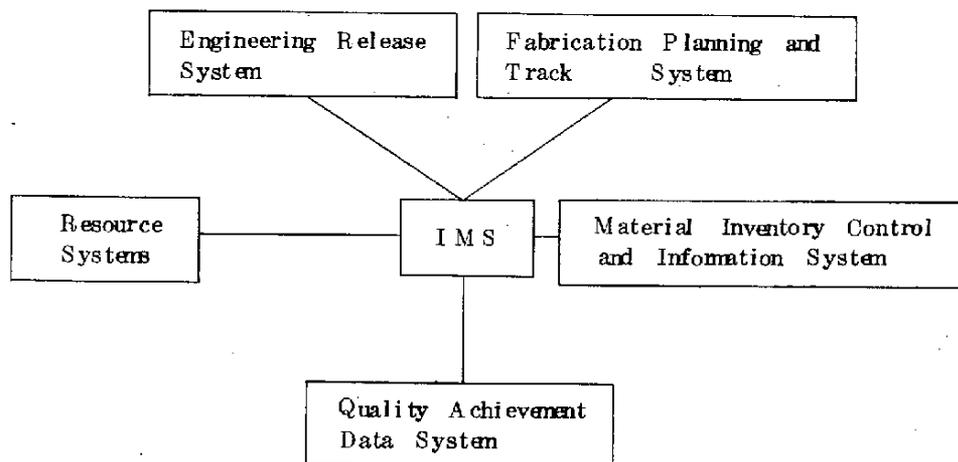


### Ⅲ 組織内データベース

#### (A) ノースアメリカン・ロックウェル社

ノースアメリカン・ロックウェルのデータベースは、1961年及び1962年にアポロのCSM（最上段の指令室）およびサターンS-IIのブースターの受注に伴う総合マネジメントシステムの中に生まれたものである。

第2.1図 ノースアメリカンのMIS



この総合システムは、製造過程における設計、仕様の管理を扱う Engineering Release System, 製造計画と、製造伝票をオンラインで管理するシステム Fabrication Planning and Tracking System, 全米にまたがって約200万に達するアポロの部品の購入、輸送、貯蔵等を管理するシステム Material Inventory Control and Information System, 高度な品質管理のためのシステム Quality Achievement Data System, 必要な人材を確保するためのシステム Resource Systems などで構成される。

そしてこれらシステムの運営に必要なほう大な情報処理のため、IBMと協同で開発したのがIMSと称される総合ファイルと多数のターミナルから構成されたオンラインのInformation System である。

#### IMSの概要

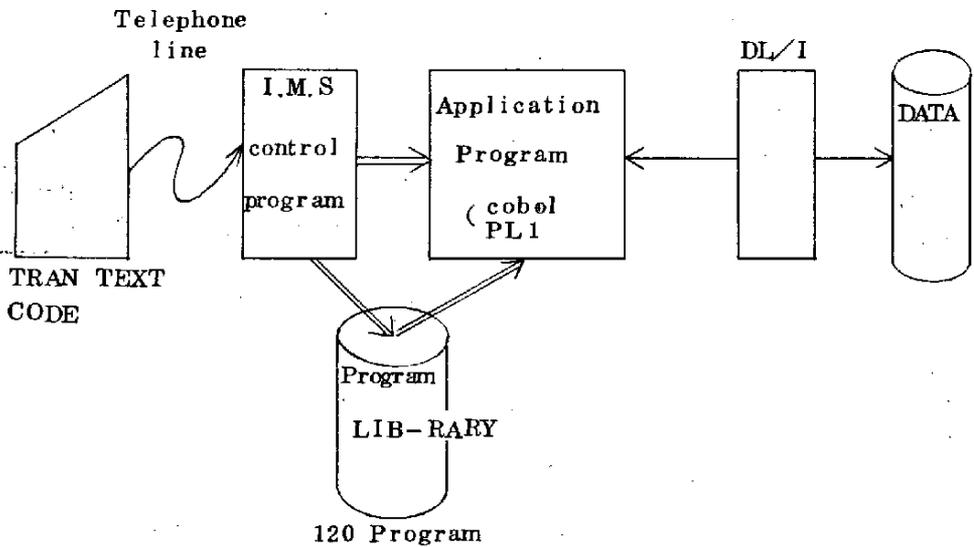
IMSは1966年より約4年間に、多方面にわたる技術が蓄積され、とくに多くのUti-

ility Program が開発されて、IMSの効率を著しく高めるのに役立っている。

Utility Program としてSpace Divisionで約120のProgramを使用しているが、その典型的な1つはERS (Engineering Release System) である。

ERSは、技術情報やデータをコントロールするもので、設計、購入、製作、組立、試験および検査等のデータを取扱っている。

第2.2図 IMSの内部構造



#### DL/Iの例

GU	GET HOLD UNIQUE
GN	GET NEXT
GNP	GET NEXT WITHIN PARENT
RPLC	REPLACE
ISRT	INSERT
DLET	DELETE

#### コンピュータシステムの概要

IMSに使用するコンピュータシステムの主計算機は、最近360/65×2より360/85に切り換えが完了し、現在、バックアップとしては、Extra Core Boxのみという

換約ぶりである。

図で示すように、360/85は、core 2,000,000 Bytesをもち、32Kのhigh speed buffer が付属している。

Direct Access Storageとしては、2ケの2301ドラムと9ケの2314ディスクをもつ。

通信線のセレクターとして6ケの2860と1ケの2870multiplexorを備え、能力としては、112回線の電話、一つの7770、audio response、および95のターミナル(2740, 1050)を処理することができる。

ターミナルは、主として、カリフォルニア、フロリダ、ミシシッピ、テキサスおよびオクラホマ州に置かれている。

3台の2701通信制御器は、6台の2848コントローラーを通じて40の2260 CRTターミナルに接続している。

## (B) ロッキード・ミサイル&スペース・カンパニー

(1) Lockheed Aircraft Corporationは約9万6千人を擁し9のCompany Division に分割されている。

Lockheed Missiles & Space Companyは、この9社のうちの1つであり、2万3千人を有し、広い分野にわたるシステム・エンジニアリング、研究開発を行なっている。当社は単なる製造業者でもなければ、単なるコンサルタントでもなく、システムのデザインから完成・操作まで一貫して取扱うシステムのディベロッパーの役割を果している。

Information Sciences Laboratory は、情報処理に関する研究開発を担当しており、エレクトロニクス、インフォメーション・リトリバル、ライブラリーサイエンス、言語学などの分野を対象として、interactive information retrieval System, data structure, などを取扱っている。

(2) 取扱う問題の内容は、下記の3種に大別できる。

### a. インフォメーション・リトリバル (情報検索)

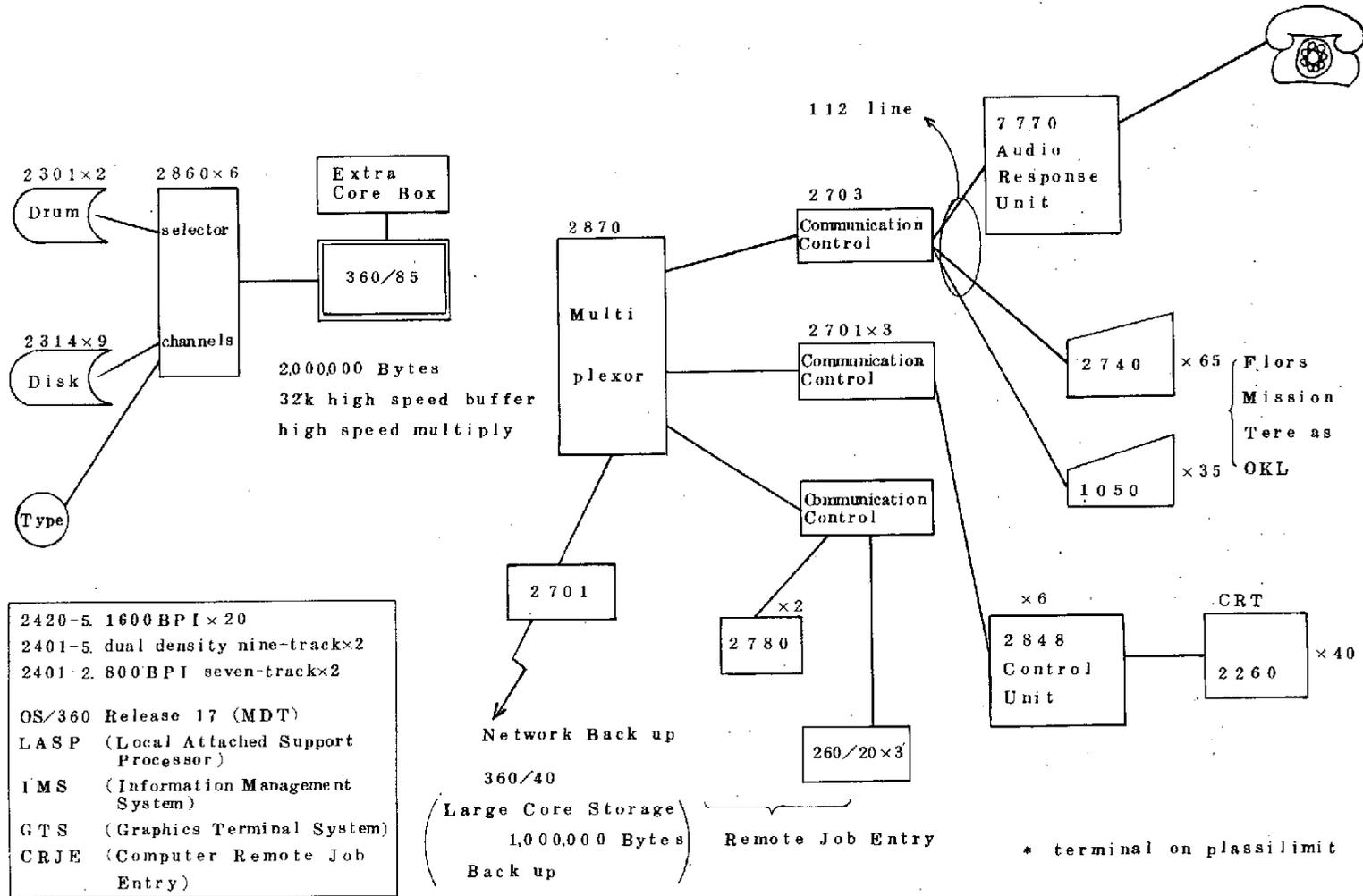
management reporting and control systems library retrieval systems などを含むインフォメーション・システムとモデルの研究開発。

### b. Interactive Systems (対話形システム)

コンピュータと人間の会話を可能にする手段の研究開発で、プログラミング・システム、

第 2.3 图 North American Rockwell Computer System

-86-



2420-5. 1600BPIx20  
2401-5. dual density nine-trackx2  
2401-2. 800BPI seven-trackx2

OS/360 Release 17 (MDT)  
LASP (Local Attached Support Processor)  
IMS (Information Management System)  
GTS (Graphics Terminal System)  
CRJE (Computer Remote Job Entry)

360/40  
(Large Core Storage)  
1,000,000 Bytes  
Back up

ランゲージ、グラフ、オンライン化問題などを取扱っている。

### c. Computing Systems

コンピュータ組織論、計算システム分析のための手法などの研究であり、ハードウェアの研究であり、ハードウェアの関係、スイッチングなどを対象としている。

- (3) インフォメーション・リトリバルに関する当社の代表的な開発例は、1966年より始められたオンライン・インフォメーション・リトリバル用のランゲージでDIALOGと呼ばれるものがある。これは、NASAに採用され、30万件のテクニカル・レポートの要約をPalo Altoにある当社のIBM360/40にストアし、NASAの各所に設置された端末機から使用されている。当初はNASAのカリフォルニア州マウンテン・ビューのAmes Reserch Centerに端末機を設置して実験を始めたが、1968年7月にワシントンのNASA、本部にも端末機を重要拡張し、以後、毎日、使用されている。
- (4) 1968年には、NASAの広範囲なインフォメーション・リトリバル・システムの開発に先がけ、Maryland a Collage ParkにあるNASAのコンピューターに収められたデータベースを全米にまたがる端末機から利用するRECONというシステムを完成したコンピューターは、IBM360/50で、リアルタイム、オンライン、タイムシェアリングによる端末機からのインフォメーション・リトリバルとなっている。このRECONシステムにより、宇宙科学文献のリトリバルが可能となっているが、限定された特定の端末機からは、コンピューターにストアされているNASAの機密文書を検索することが出来る。

### (C) ニューヨーク連邦準備銀行

ニューヨーク連銀は、アメリカの中央銀行組織たる連邦準備制度を構成する全米12の連邦準備の一つであって、ニューヨーク州の全部およびコネティカット、ニュージャージー州の一部を管轄し、銀行券の発行、加盟銀行に対する準備金の保持、手形割引、有価証券売買ならびに政府の金融業務代行などの中央銀行としての機能を遂行している。

コンピュータ利用は、調査統計局の経済統計データバンクに関するもののほかそれぞれの固有業務のために行なっているが、われわれは調査統計局のデータバンクに限定して述べてみる。

#### (1) 経済統計データバンクの現況

約2年前から、ニューヨーク市内にある各種機関の調査スタッフとともに"Project Economics"と称する非公式の組織を結成し、TSSによる経済統計データバンクのシステム開発を進めつつある。

現在、このProject Economicsに参加しているメンバーの数は約20であるが、

参加を希望して待機中のメンバーはその倍に及んでいる。収録データは国民所得統計をはじめとする一般経済指標の時系列データで、系列数は約550、これに要する記憶容量は5,000,000ワードに達しており、当面の計画完成時には、4,000,000ワードが必要となる見込である。550系列の内容は第1表の通りである。現在使用中の装置は98KBのMemory Module 3セット、6.5KBのディスク1セット、MT装置16セット、データ通信回線16ライン、CRTディスプレイその他の端末装置などである。

ソフトウェアの問題は、この種のシステムをうまく運営してゆく上で最も重要なポイントであるが、この点について連銀はいま独自のシステム開発を進めつつあるところである。

上のごとく、NY連銀のデータバンクは現在なお、開発中の段階である。

データバンクの利用方式もいわばクラブ制でコーザーの範囲もきわめて限られている。今後の見通しとしても連銀は、この種のデータバンク・サービスを一般利用希望者に対して提供することは全く考えていないことを明らかにしている。

第1表 データバンクの収録項目

項	目	系	列	数
(国民所得統計)				
A	国民総生産			
B	企業在庫増減			
C	外国部門			
D	政府部門			
E	GNP 主要生産タイプ別			
F	国民所得(業種別企業利潤をふくむ)			
G	個人所得			
H	政府バランス			
I (金融財政統計)				
	商業銀行			
	金融市場レート			
	公債、株式利回り			
	株式市場データ			
	消費者信用			
	不動産信用			

項	目	系	列	数
	住宅貸付残高			
	(その他)			
J	企業倒産			
K	資金収支			
L	製造業者出荷高			
M	小売業売上高			
N	製造業者新規受注および受注残			
O	在庫			
P	物価指数			
	卸売物価			
	消費者物価			
Q	固定資本形成(建設業をふくむ)			
R	企業の設備投資支出			
S	企業コスト(比率)			
T	商業銀行の資産負債			
U	ニューヨーク市中銀行の資産負債			
V	米国国際収支			
W	鉱工業生産指数			
X	人口、労働力、雇用、家計			
Y	米国政府予算(財政)			
Z	分析的比率(稼働率他)			

(2) データバンクの定義ならびにシステム開発に関するNY連銀の見解

1) 用語の定義

○ Information Utility (情報の公共利用)

「マス・ストレージを有する1台のコンピュータが、数多くのユーザーとの間にオンラインでつながれている状態」を指す。

○ Data Bank

「多数のユーザーがアクセス可能な集中化されたファイル」を意味し、特定の関連をもつデータの集積であるが、通常、相互に重複しないデータ・セットの集まりとして構

成される。

## ○ Data Base

「1つの企業または機関のCentralized File」のこと。したがって、オンライン、オフラインのデータ・セットから成るデータ集積の全体であり、また、1つまたは多数のData Bank, Data Bank に入らないデータ・セットを包含する。

## 2) データ・ベース・システムの開発

データ・ベース・システムを構成する2大要素は、データ・ファイルとユーザーであって、システム設計の目的はこのファイルとユーザー間のコミュニケーションを可能ならしめることにある。この目的を達するためには人、ハードウェア、ソフトウェア三者の組合せを効率的に行なうことが必要であるが、データ・ベースとユーザー間のコミュニケーションを実現して行く上で特に重要なポイントとして次のものがあげられる。

### ① Operating System

OSとはコンピュータの各構成部分におけるオペレーションや、プログラム相互間のコントロールとスケジューリングのためのプログラム・システムを包括するのであって、データ・バンク・システムが実用化されるにつれて、より高度なOSの開発がますます重要視されるようになった。

### ② Data Management Language

言語の面では、プログラミングを知らない分析者でも直接アクセスできるよう、日常用語に近いものの開発が要望されるようになってきた。当連銀でも、そういった趨勢に即して独自の"Research Access Language"を開発する計画をもっている。

### ③ Application Program およびアセンブラー、コンパイラ

レポート作成のためのアプリケーション・プログラムならびにそのための言語の開発も、データベース・システムの効率的な運営のためには必要である。NY連銀では、COBOL を使ってコマンド方式によるアプリケーションを開発した。

### ④ ファイル構成

ユーザーがファイルに直接アクセスする場合のファイル構成の方式には(i)Sequential (ii)Chain(List) (iii)Indexed Sequential (iv)Hierarchical (v)Randomなど、各種のものがあ、ユーザー側のデータ保存、再新検索上の必要性に応じて適切な構成方法が採用される。経済指標のデータバンクでは、これらのうち(iii) (iv)あるいは両者の組合せが、更新ならびに検索上有効なことが多い。

### ⑤ ユーザーのニーズに対応するための配慮

「データベース・システムに対するユーザーのニーズとして考慮すべき重要なファクターは(i)システムのレスポンス・タイム (ii)利用のタイミングおよび頻度があらかじめわかっているかどうかの2つである。」

データベース・システムに関する問題としては以上の他、ファイルの内容を表示する形式、ファイルの秘密保持、将来の変更に対する配慮などがある。

現実にデータベース・システムを開発、設計してゆくにあたっては、それぞれの置かれた環境に応じて試行錯誤をくり返して行くことがどうしても必要になる。上述は、ごく限られた側面に関する単純な説明にすぎない。

#### (D) ボストン小児科病院

##### ○ システムの目的

病院は航空機やホテルと同じように、数に限りのある場所を持ち、常に人々が出たり入りたりしている。場所を最大限に利用する点については、問題は共通しているわけである。自動予約システムが航空機やホテルの目的を達成するのを助けてくれるのであれば、これは病院の同じ目的にも役立つはずである。

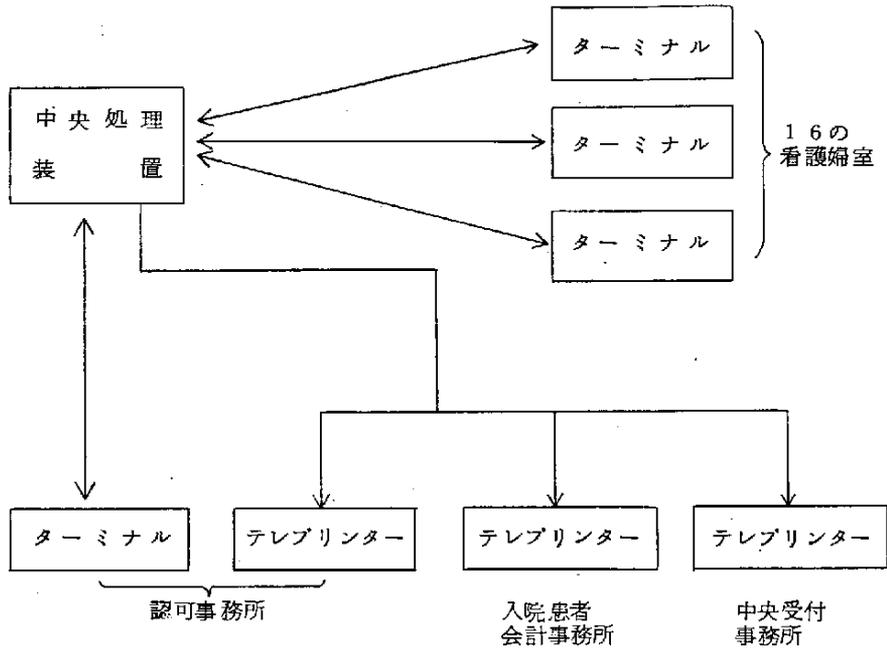
ボストンの小児科病院医療センターの350のベッドにつき、自動患者所在コントロール・システムを企画したのは、この理由のためであった。この場合、経営コンサルタントのCresap, McComick及びPagatの助けを借りた。このシステムは2つの試みの後、現在稼動しているがピークの占有度を取扱うことを助ける主目的を立派に達成するばかりでなく、副産物としていくつかの価値のある恩恵をもたらすようになった。院内の連絡を改善したこと、認可手続を早くしたこと、患者の状態と所在に関する最新情報を提供できるなどの恩恵は決して小さいものでない。

##### a) 稼動中のコントロール・システム

患者の所在コントロール・システムは、基本的には4つの要素から成立する。

- (1) 認可事務所における入出力のターミナル
- (2) 患者情報を蓄積、索引、報告するコンピュータ
- (3) 16の看護婦室の入出力のターミナル
- (4) 認可、会計、受付の3つの事務所におけるテレタイプ・プリンター

第2.4図 ポスト小児科病院医療センターの患者所有コントロール・システムの構成



患者が病院に来ると、事務員が患者の名前、生年月日、男女別、治療記録番号、電話番号、認可番号、認可月日、「科」などをタイプする。この情報の複写が、患者の前のブラウン管に写して確認した後、事務員はランダム・アクセスのできるディスクファイルに蓄積されるコンピュータにこの情報を送るため、ボタンを押す。

一応ベッドが割当てられると、その後の患者に関するコンピュータとの連絡は、ベッド番号によってなされる。

b) 患者の看護データ

ベッドが割当てられると、看護婦は患者の状態、他科への移転、退院見込、死亡、退院その他の情報を報告する。彼女はその報告が正確かどうかスクリーンで確かめる。

このメッセージの最大の利益は、情報センターを所有できることである。

患者情報の正確さと利用度の高いことが、電話呼び出しを縮少し、看護婦室の書類を減じた。

看護婦室のターミナルは、小さなブラウン管表示装置と数個のキーをもっている。キーボードはベッドの確認に使われる10の数字によるキーボードである。

このシステムは現在1日に少なくとも600のメッセージを伝えているが、誤ったメッセージは唯一つにすぎない。

c) 達成した成果

自動コントロールシステムによって、タイムラックは完全に除去された。現在病院はリアルタイムで操業されている。したがって、すべての入院患者が何処にいるか、彼等の医療状態はどうなっているか等が1分内に正確に知らされる。

ベッドの最適利用は、このインフォメーション・システムの最も重要な成果である。これは小児科病院にとって、役立つ情報である。これによって、患者を早産児から青年までの大きさの順に並べたり、性別に分けるのに役立つ。

d) トータル・システムに向って

患者所在コントロール・システムは最終目標に向っての第一歩である。その最終目標は病院の医療、統計、会計データのすべてを、単一の情報システムに統合することである。次頁図は、トータル・システムがコンピュータと病院内の重要箇所との間にオンラインのコミュニケーションが如何に進展しているかを示すものである。

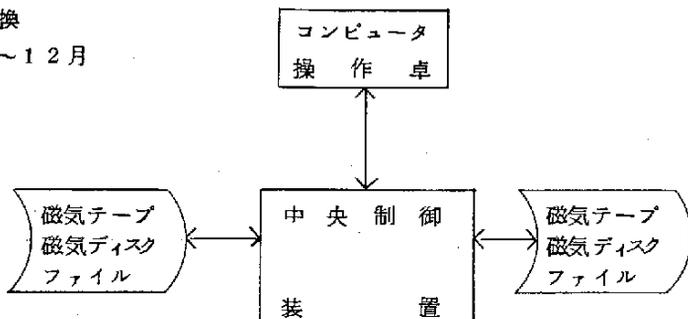
3段階の実施計画では、トータル・システムが1968年の終りに、動くようになるとみている。

第2.5図 小児科病院医療センターの病院情報システムの構成

第1段階

財政への適当の転換

1965年7月~12月



第2段階

コンピュータへのオン

・ラインによるコミュニケーションの開発設計

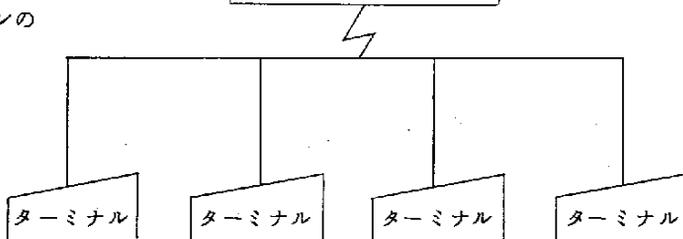
1966年1月~12月



第3段階

ターミナルへのデータ  
コミュニケーションの  
実施

入出力ターミナルから  
データ伝送のコントロール



テレタイプ、タイプライター  
ブラウン管表示装置

- | 認可       | 医療記録     | 補助部門      | 看護婦室     |
|----------|----------|-----------|----------|
| ○ ベッド利用  | ○ 所在     | ○ 実験室のテスト | ○ 患者の看護  |
| (患者所在)   | ○ 患者ファイル | ○ 検査      | ○ 補助サービス |
| (コントロール) | ○ 調査の利用  | ○ 結果      | ○ ベッド利用  |
| ○ 診断計画   |          |           |          |

## (E) ペンシルヴァニア大学

ウォートン、エコノメトリック・フォーキャスティング・アンシエツ

- (1) ペンシルヴァニア大学のビジネス・スクールであるウォートン・スクールにおいては、エコノメトリックのクライン教授を中心として10年にわたるアメリカ経済予測モデルの研究開発が進められてきた。1963年にEconometric and Forecasting Unitの組織作りが行なわれ、これが発展してペンシルヴァニア大学の非営利組織として、Wharton Econometric Forecasting Associates Inc. となったものである。
- (2) 当研究所は、経済予測とエコノメトリックスの研究を巾広く行なうとともに各種のプロジェクトの共同作業も進めている。また、現在ではメキシコなど海外予測モデル開発などのプロジェクトも持っている。
- (3) 活動内容の主たるものは
  - a. アメリカ経済のエコノメトリック・モデルによる予測
  - b. 産業連関表によるアメリカ50部門の産業予測
  - c. エコノメトリックスなどの法の基礎研究、金融モデルなど新分野でのモデル開発、コンピュータ技術の研究開発などの研究活動
  - d. セミナー、予測会議などの教育活動がある。
- (4) データバンクとしては、Data Resources, Inc. との提携の他にブルッキング研究所のデータバンク、LINKモデルのデータバンク、当所研究所のデータバンクなど多角的な連携がある。
- (5) コンピュータ・システムとしては、IBM360/75を有するコンピュータ・センターとリモート・ターミナルで結びタイムシェアリングしている。

アメリカ経済予測モデルはフルメンバーとして米国主要企業21社との共同予測作業であり、雑誌「ビジネス・ウィーク」に発表している。

またメキシコ・モデルはメキシコの15社との共同作業であるなど、当研究所の予測作業は実務家との深いつながりのうえに運営されているという特徴を有している。

## (F) 日本開発銀行

開銀が現在使用しているコンピュータはFACOM230-50およびFACOM230-30各1機で、適用分野は大別して、①日常業務、②経営分析、③調査企画、の3業務に分けられる。作業の性格という面からみれば、①データの記録、②データを利用したの定期的計表

作成、随時抽出、③経営科学手法の活用、の3つに分けることもできる。

#### (1) 日常業務

本行融資先の取引状況に関するデータを収録、集中管理して、それをもとに各種業務資料を作成、提供するものである。この面での機械化はすでにPOS時代から行なわれていたが、コンピュータによってさらに体系化が進んだ。そのほかに、各職員の給与、経歴に関するデータも収録し、人事資料を作成している。

#### (2) 経営分析

財務諸表データは企業別情報の基本的なもので、利用範囲が広く、原資料の収集、体系化も比較的容易である。昭和39年からコンピュータ収録に着手し、有価証券報告書ベースによる上場会社のデータ蓄積は、すでに第1部、第2部の大半を完了し、速報性には之しいが、データの豊富性、信頼性においてすぐれていることから、銀行内では最も広く利用されている。

企業の経営計画や融資対象プロジェクトを長期にわたって綿密に検討するため、計算方式や因果関係を体系化して、外生値さえ与えれば、幾通りでも計算比較できるように、現在、長期計画検討用6種、個別プロジェクト4種のプログラムがあり、貸付、審査等の際に利用されている。

#### (3) 調査計画

アンケート方式による調査を3種類、(イ)設備投資計画調査、(ロ)地域設備投資計画ならびに融資期待額調査、(ハ)工場分散調査を行ない、その集計作業をコンピュータで処理している。経済産業統計のうち、主要なものを長期間にわたって組織的体系的に蓄積し、これを分析の必要に応じて機動的に抽出加工できるようにしている。現在、主要経済統計として、国民所得統計、法人企業統計、産業設備資金統計を収録しているが、これらは情報交換の形でインプット・データシートを直接所管省庁から入手している。そのほかに、地域経済統計、企業別業界統計も収録している。

#### (4) K E M P F

開銀は、39年度末機種決定後直ちに富士通との協力のもと、「汎用的プログラム・システム」の開発に着手し、41年7月導入時まで当初考えていたシステムの大半を完成し、これをK E M P Fと名づけた。K E M P Fは計量、モデル分析、O R手法、多変量解析、統計解析のほかデータの収録、抽出、加工機能もふくむ。

#### (5) データ・バンク利用上の問題点

開銀が業務遂行上取扱う多種多様の情報のほとんどは外部からの入手に依存している。し

かも、政府の政策金融機関という性格上、経済社会の変化に即応した機動的、有機的な情報処理活動が要求される。このような立場からコンピュータ・システムと情報活動には次のような問題点がある。

#### 1) 情報の利用側、提供側のギャップとその調整

人間の扱う情報は、その選択が個別的であり、しかもニーズは時とともに変化する。提供側の悩みはこういった選択の個別性とニーズの変化に適確に対処できるか、また利用者の不時の要請にいつでも応じ得るだけの大量のデータを常にかかえこむことができるかということである。

この問題を根本的に解決するには、タイム・シェアリングの実現によってコンピュータと通信回線の結合による情報処理の迅速化にまっぴらではないであろう。次善的な解決策としては、内部的には利用者と提供者の接近をはかり、対外的には各関係機関相互間に情報処理ネットワークを設けることが考えられている。

#### 2) 組織上の問題

利用側は依頼者、提供側は受注者というような形で責任を分担するのではなく、両者がサブ・システムあるいはコンポネンとして1つの組織を形成することが望ましい。社内的には固定した部を撤廃して機能別集団とし、プロジェクトごとに各集団から参加するチームを編成すれば、情報処理がスムーズに生かされるのである。

#### 3) 外部関係機関への要望

行政機関の作成する統計は、戦後の統計整備の必要性の認識を基調として発展してきたもので、20年以上経った現在では指定統計の体系を抜本的、飛躍的に改善、整備することが望ましい。とくに、生産力重視の観点から、農林、鉱工業、運輸、労働力等の関係統計が不必要に膨大している反面、社会変動に関する統計が著しく立ちおけていることは問題である。

民間機関では、設備投資アンケート調査や財務情報などの供給面で重複の面が見られる。全体として効率的合理的なデータ・バンク利用システムの開発が望まれる。データ・バンクの開発・利用は、特色をもった専門機関が多元的に存在し、相互間で情報交換可能というネットワーク構造を前提として進められることが望ましい。

### (G) 日本長期信用銀行

長銀は普通の市中銀行とは違って、融資はいくら短くても5年、電力融資の場合には15年に及ぶ。したがって融資に際してできるだけ会社の実状をくわしく洗おうとする。その場合企

業の実態を現わす指表はやはり財務諸表である。それを少なくとも過去3年間位は調べる必要がある。

以前は担当者がバランス・シートなり損益計算書、あるいは有価証券報告書などを貰って調べていた。おかしいと思えば組み変えて修正バランスを作製したりしていたが、日数がかかりすぎた。

業種別には、平均的に一体どうなのか、あるいは大手と中小とではどうなのかということから、28、9年から資料を集めた。また、それに対してランク付けをするのはどうかということになり、そのために一番手掛りとなるのは財務諸表だという結論に達した。その後、コンピュータを使用して出発することになった。データそのものは昔から持っていたという理由もある。

データの種類もコンピュータに入れる以上、画一的なものでなければならないということを考えると営業報告書では困る。そこで議論に議論を重ねられ、有価証券報告書が採用された。それがないところは、やむをえぬということになった。

以上のように目的とニーズは内部から持ち上って来た。この施設をオープンにするという計画もある。以上の様な形成過程は図1に示す通りである。

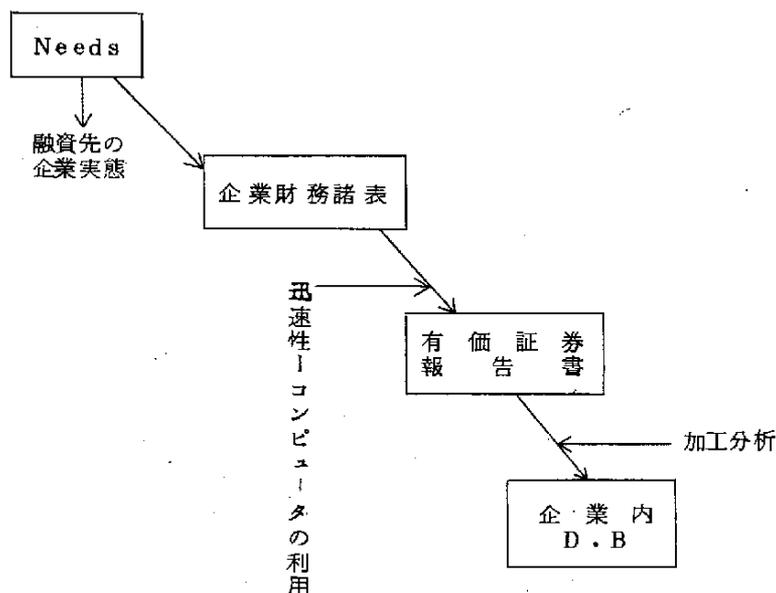
長銀はデータをoutputして原点へフィードバックしており、原点が借出しをする場合にそれを参考にしている。

また本部ではA社とB社を比較するのに使用したりしている。現在、作業しているのは、バランスの分析と資金移動、それから他銀行からの借入状況それから財務データに比率をつけ(重味付け)業種別に比較してみることなどである。

又、年々の資金状況調査を年に2~3回やり、関連性をみるなどということも考えている。

長銀が一番苦勞するのはinputのデータ集めである。取り引き先から借りてくるものはよいが、必ずしも全部揃うとは限らない。これは非常に労力を要することであり、そこで約200項目にしぼってデータを整理している。プログラムも二年掛けて相当長いものだが一応完成した。

第2.7図 長銀のD・B形成過程



## (H) 関西電力

### 1. 電力産業の情報ニーズ

電気エネルギーは光速に近い早さで流通し、また生産即消費であり、在庫が実質的に零である。地域独占の形態をとり広大な面積に供給しなければならないので、普通の企業とは異なった情報を必要とする。

生産・流通を管理する情報は、その速さの点で次の3つのシステムに分類される。

#### 1) 高速管理システム

落雷のような不測の事故を迅速に除去するために使われる。関西電力では、事故の発見および必要箇所への伝達速度は約50分の1秒から50分の20秒、そして受けた情報にもついて実際に電気を切る早さは50分の20秒から1秒の間で実施しなければならない。

#### 2) 中速管理システム

高速管理システムで除去できなかった事故を処理するために、2秒から数分の間で情報処理が行なわれる。このシステムは現在もっとも遅れている。電話では遅すぎるし、所要情報ルートが複雑で利用度も低いからである。1965年のニューヨークの大停電はこの

中速管理システムの欠陥によって生じた。

### 3) 低速管理システム

発電機の出力行コントロール、系統の切替え等、数分から十数分の速さで処理すればよい情報については現在電話で処理されている。しかし、最近では、この部門にもコンピュータの進出が目立ち、近い将来には完全にコンピュータライズされていく可能性がある。

経理、資材、労務関係の情報は一般産業と類似しているが、販売に必要な情報は、重要家数が多く、広く分布し、しかも変動するという点でデータ処理および業務管理上特殊な問題を持っている。また、電力設備をつくるのには約5年かかるので、5年後の設備計画、収支予測等の情報が要請される。したがって、10年とか20年とかの非常に長期のシミュレーション分析が必要になる。この場合、情報速度よりも内容の深さが重要になる。

一般情報ネットワークとのアナロジーで、流通系統は従来信頼度を高めるため、コストの許すかぎりmeshが選択されてきたが、最近系統の簡明さ、大規模事故(たとえばニューヨーク大停電)の復旧の早さ等からTreeシステムが再認識されつつある。

## 2. 関西電力の情報処理システム

関西電力は、1967年7月に、IBMシステム360のモデル40と50をコンバインした形で導入している。ELD運用は、実際にはIBM360/50で主な計算をやり、その結果をIBM1205によって1200ボアの通信回線で10kmぐらい離れた本社へ送る。本社ではHITAC8300がその情報を受け、必要な箇所、たとえば発電所はかなり低速で指令を発する。必要なインフォメーションはテレメータで本社へ集めてHITACが整理し、IBMに廻すというデータ・システムである。

おもしろい使い方としては、オンライン・インクワイアリーがかなり順調にしていることである。2kmから60km離れた9営業所からIBM1050を使って行なっているが、1968年11月から稼動して対象軒数は約200万軒、毎月2万ないし3万のインクワイアリーがある。

料金調定のために、約550万需要家を対象に過去3ヶ月の調定実績異動事項等を含むマスターテープをつくり、検針票作成、料金計算、領収証作成等を行なう。このぼう大な処理には1日27万軒の読みとりができるOCRを使っているが、実際のデータを送るにはトラック輸送によっている。

### 1) 長期経営分析ファイル

いくつか持っている長期経営分析のシミュレーション・モデルを管理するファイルをつくりたい。シミュレーション・モデルは計算するたびに修正を加えるので、修正内容を管

理するファイルと、その上に各種の初期条件、パラメータ値、テーブル値などのシミュレーション条件を適切に管理する必要がある。そうしておけば、もしこういう条件で計算しろというとき、その条件だけを計算するというシステムになる。それから、シミュレーションの主な計算結果もファイルしておきたい。

これらの構成は、データファイルとしては3種類ぐらいの内容を持つもので、それがGISによって管理され、インクワイアリーは直接GISからきくことになる。

## (1) 野村総合研究所

当社で財務データ・ファイルと呼んでいるものは、バンクというようなものではなく、財務諸表を考えたものである。それに必要なのは、網羅性(一部上場会社中心)、連続性、迅速性の三つである。

用途はプロパーで当面使うことから出発し、考え方はいろいろあるが、

- ① プロパーで使う
- ② 一部オープンにする (整備されていて、要求があった場合)
- ③ 最初からオープンを目的とする (データ蓄積と同時に特定のユーザーの為にIN-PUTすることもある)

という三つがあり、当社はプロパー使用の方針をたて、その使用目的に、

- ① セミ・マクロに産業界の動向を迅速にとらえる。
- ② 個々の産業界のデータをマイクロベースでとらえる。
- ③ 株価との関連を調査する。

この三点を掲げた。

政府機関との関連については、データ・バンクは時間的な要素を含めてPaylineに乗る必要性から、政府機関が行なうものは自づから範囲が決まってくると考えられる。

ファイルの概略

### ① 財務諸表のファイル

東商第一部上場640社-39年以降報告書と有価証券報告書を調べ、55項目

input 204項目Outputしている。

主要300社-30年以降税務報告書をベースにinput

(39年商法改正があったが、30年まで遡り、修正してinput)

- ② 業積予想-東商第一部上場社について、実績と予想を簡単な別項目に決め、これによってランキングをつける。

③ 東商第一部の39年以降の株価、投資その他の資本移動の項目を勘案したもので、今問題になっている一株当りの指表が出ているが、これを全てダウ修正する。

データ処理業務について、プロパーの領域については、不十分であり、データ処理に関連した研究調査の業務の根本的相異からも考慮の余地があるという段階である。

以上であるが、網羅性、連続性、迅速性という三つの要求は現在でも変わらない。

## (J) 鉄鋼連盟

鉄鋼連盟は下記のようなデータバンク構想を持つ。

### 1. 鉄鋼情報システム確立の必要性

#### (1) 鉄鋼業界における情報の利用

##### a. プランニング情報とオペレーティング情報

情報を機能的な視点からとらえると、意思決定のためのプランニングに用いられるプランニング情報と日常業務の処理に関するオペレーティング情報に分けられる。前者はトップ・マネジメントおよびミドル・マネジメントのための情報であり、後者は現業、部門の情報と言える。

##### b. 社外情報と社内情報

情報を現象的な視点からとらえると会員会社においては、社外情報と社内情報とがある。

鉄鋼情報システムで扱われる情報は、プランニング情報であり、会員会社から見て社外情報である。ここで重要な点は、業界のプランニング機能には業界外の外部情報が不可欠であり、また会員会社におけるプランニング機能には社外情報（業界内情報と業界外の外部情報）が不可欠なことである。

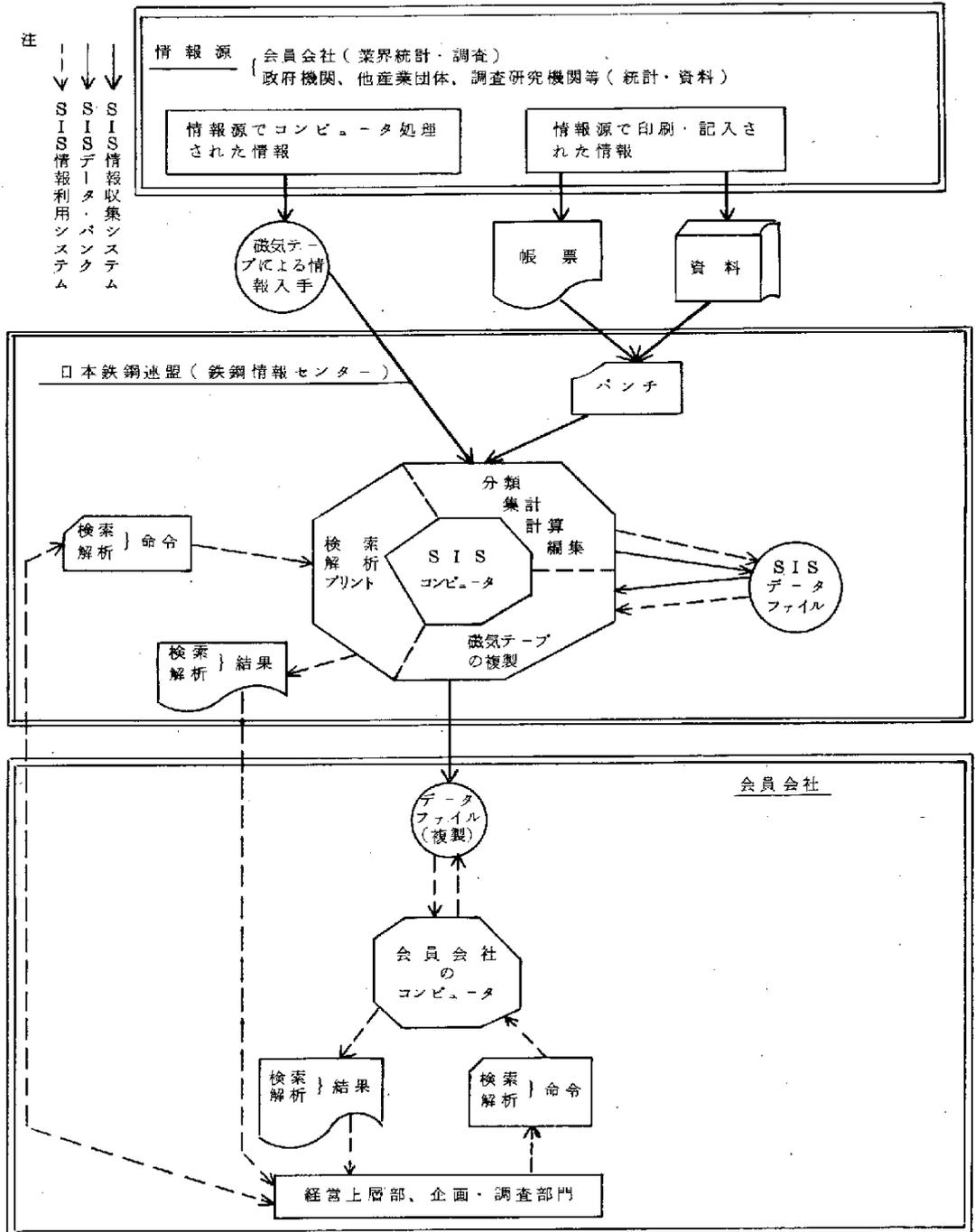
鉄鋼情報システムは、これらの情報を体系的に整備・蓄積することを中心に、会員会社への情報提供サービスの方法をも含めた情報の利用の仕方、このシステムを効率よく収集、入手する方法をも含めた総合的なシステムとして設計されるものである。

#### (2) 経営情報の蓄積

鉄鋼情報システムは、日本鉄鋼連盟が鉄鋼業界を代表して入手し、収集している情報のうち、業界ベースでも会員会社ベースでもプランニング情報として活用できるものを利用のニーズに基づいてコンピュータに蓄積し、各会員会社が必要な情報の磁気テープにコピーを提供することによって多くの社外情報が蓄積できるようにしようという構想である。

(3) 鉄鋼情報システムの概要

第2・9図 鉄鋼情報システム概要図



#### (4) 会員会社におけるメリット

鉄鋼情報システムが会員会社にもたらすメリットとしては次のような効果が考えられる。

- a. 社外情報の蓄積の容易化
- b. プランニング情報の高度利用
- c. 情報処理における省力化

## 2. 鉄鋼情報システムの構成および内容

### (1) 鉄鋼情報システムの全体構造

鉄鋼情報システム(SIS)の全体の構造は、SIS情報収集システム、SISデータ・バンク、SIS情報利用システムの3つのサブ・システムで構成されることとなる。

- a. SIS情報収集システム — 情報入手のためのシステム
- b. SISデータ・バンク — 情報の蓄積機能のためのシステム
- c. SIS情報利用システム — 情報の利用機能のためのシステム

このうち、SIS情報収集システムおよびSISデータ・バンクのかんりの部分と、SIS情報利用システムの全部とにコンピュータを用いることが、鉄鋼情報システムの効率を高めると考えられる。

### (2) 各サブ・システムの内容とシステム化の要点

#### ① SIS情報収集システム1：統計調査システム

このサブ・システムの検討の要点は、両方の処理手順のフローをつなぎ合わせて改善を検討することで、コンピュータに直接インプットすることができる磁気テープによってデータを提出してもらうことが一つの方法として考えられる。これによって会員各社および日本鉄鋼連盟事務局の両方で省力化(工数の節約、人員の節約)でき、また集計結果作成の迅速化を図ることが可能である。

#### ② SIS情報収集システム2：資料情報収集システム

統計データ等の先方でコンピュータによる処理が行なわれている部分については、先方と協議の上情報交換ネットワークを形成して、先方で処理した磁気テープのコピーを入手する。これにより現在の調査業務における情報の入手、処理、蓄積等の活動における省力化、能率化、迅速化のメリットが期待できる。もう一つの要点は、新たに必要となった情報を入手し、蓄積できる体制の確立と、特に従来手薄だった海外情報の入手、収集体制の確立である。

#### ③ SISデータ・バンク1：汎用統計データ・システム

設計されたファイルの様式に従って原データと集計結果とをコンピュータによって編集し、磁気テープに記録して保存するところにこのサブ・システムの検討の要点があり、このように蓄積されたデータの検索、解析、利用をコンピュータで行なう際にデータを直接インプットすることができる。

④ SISデータ・バンク2：目的別統計データ・システム

鉄鋼業界におけるプランニング機能には、種々の情報源から多くの項目の統計データを入手して分類編集してファイルしておき、必要の都度タイムリーに解析作業にインプットできるようにしておかねばならないものがある。このように重要度の高いプランニング機能に用いるデータを常にアップツデーに管理しておく必要があるファイルを、汎用統計データ・ファイルおよびSIS情報収集システムにより入手した外部情報から編集して常備しておくのがこのサブ・システムであり、汎用統計ファイル同様解析利用に際して直接コンピュータにインプットできるような磁気テープを用いてデータの蓄積を行なう。

⑤ SISデータ・バンク3：分野別情報ファイルシステム

文字情報の中には、きまった分野の仕事に用いられるもので、データの項目や諸元をきめて、標準化したファイルに情報を蓄積し、常に最新の情報に更新して管理できるものがある。このサブ・システムはこのような性格の情報を磁気テープを用いてファイルするシステムである。このサブ・システムにおける検討の要点は、常にアップツデーな情報をファイルしておくために、ファイルのメンテナンスを行なう作業および検索、利用の能率化のために磁気テープを用いることにある。

⑥ SISデータ・バンク4：ドクメンテーション・システム

入手情報および作成情報には印刷物の形のレポートがある。また、新聞記事のスクラップや雑誌類の論文、文献等がこれにあたる。このような情報は、従来の資料室におけるドクメンテーションを強化して、能率的な検索システムを確立する方向で検討する。

第一次のステップとして検索利用の効率を高めるよう分類方法の改善、キーワードおよび抄録の採用等を検討しつつドクメンテーションを強化し、第二次のステップとして第一次ステップを踏み台としてマイクロ・フィルムを活用、コンピュータによる文献検索システムの開発（シソーラスの開発、利用を含むIRシステム）を行なう。

⑦ SIS情報利用システム

このサブ・システムでは、必要な情報を検索し、抽出するプログラム、情報を編集するプログラム、各種の方程式や解析モデルのプログラムを開発して、磁気テープにプロ

グラム・ライブラリーとして蓄積し、プログラムの組合せによっていかなる解析、処理もコンピュータによって行なうことができるようにする。データ・バンクおよびプログラム・ライブラリーの充実によってコンピュータの処理時間が数時間もかかる大きな解析モデルを除いて、鉄鋼情報センターに解析申込書が提出されてから申込者に結果を提供するまでの時間(ターン・アラウンド・タイム)を1～2時間とすることがこのサブ・システムの将来目標でもある。

### (3) 鉄鋼情報システムの対象となる情報

鉄鋼情報システムが扱う情報は、もちろん、鉄鋼業界すなわち会員各社および日本鉄鋼連盟事務局におけるプランニング機能に使われる情報である。その意味では、各分野における企画、調査活動のニーズから判断されるべきであり、しかも現状の問題にとどまらず、将来の問題をも見通した上での情報のニーズが認識されている必要がある。

## 3. 鉄鋼情報システムにおけるコンピュータの利用

鉄鋼情報システムの立案に際して、各サブ・システム間のインターフェースの処理を検討してみた結果、鉄鋼情報システムの運用には、コンピュータの有効利用が不可欠の条件として浮び上がってきたのである。なお、コンピュータの導入機種を選定等の導入計画は、さらに鉄鋼情報システムの具体化の為の作業を進めて、情報の処理方法および情報量などを判断できる時点で立案する考えである。

- a. コンピュータのシステム構成は、磁気テープをベースに磁気ディスクを加えたバッチ処理のシステムを前提に考える。
- b. コンピュータの稼動については、テスト・ランを除く本番稼動時間は次のような目標で計画する。

- (a) 統計集計作業
- (b) データ・ファイルのメンテナンス作業
- (c) 情報検索および解析計算
- (d) コンバージョン

なお、これによるとSISデータ・バンクのためのファイル化に要する作業(上記(a)および(b)が60～70%、SISデータ・バンクに蓄積した情報の利用、提供のための作業が30～40%という配分になる。また見方を変えて定期的に繰返し行なわれる定常作業の比率は80%～90%である。

- c. コンピュータの選定のもっとも重要な要因の一つである情報量については、現在システ

ム分析作業を開始して検討中であるが、なお、今後の作業において具体的に詳細な検討を行なう必要がある。

#### 4. 鉄鋼情報システムの運営体制

##### (1) 鉄鋼情報センターの確立

日本鉄鋼連盟事務局の主要な機能として、内部管理機能、情報処理機能、渉外機能、プランニング機能、オペレーション機能の5つの機能をあげることができる。鉄鋼情報センターは、鉄鋼情報システムの運営のセンターであり、日本鉄鋼連盟事務局の情報処理機能を分掌する唯一の職制として確立される。

##### (2) 円滑な運営のための体制

###### ① 機密に属する情報の取扱い

機密情報の取扱いについては、鉄鋼情報システムの実施に際してもそれぞれの情報の公表範囲(利用者の対象)を変える等の変更は一切行なわない。また、システム自体に機密保持を保證する機能を持たせるよう、各データ・ファイルに利用可能な範囲を示すコードを入れて、検索命令における利用者コードとつき合わせる等の方法を探ることも可能である。

###### ② データ・コードの標準化

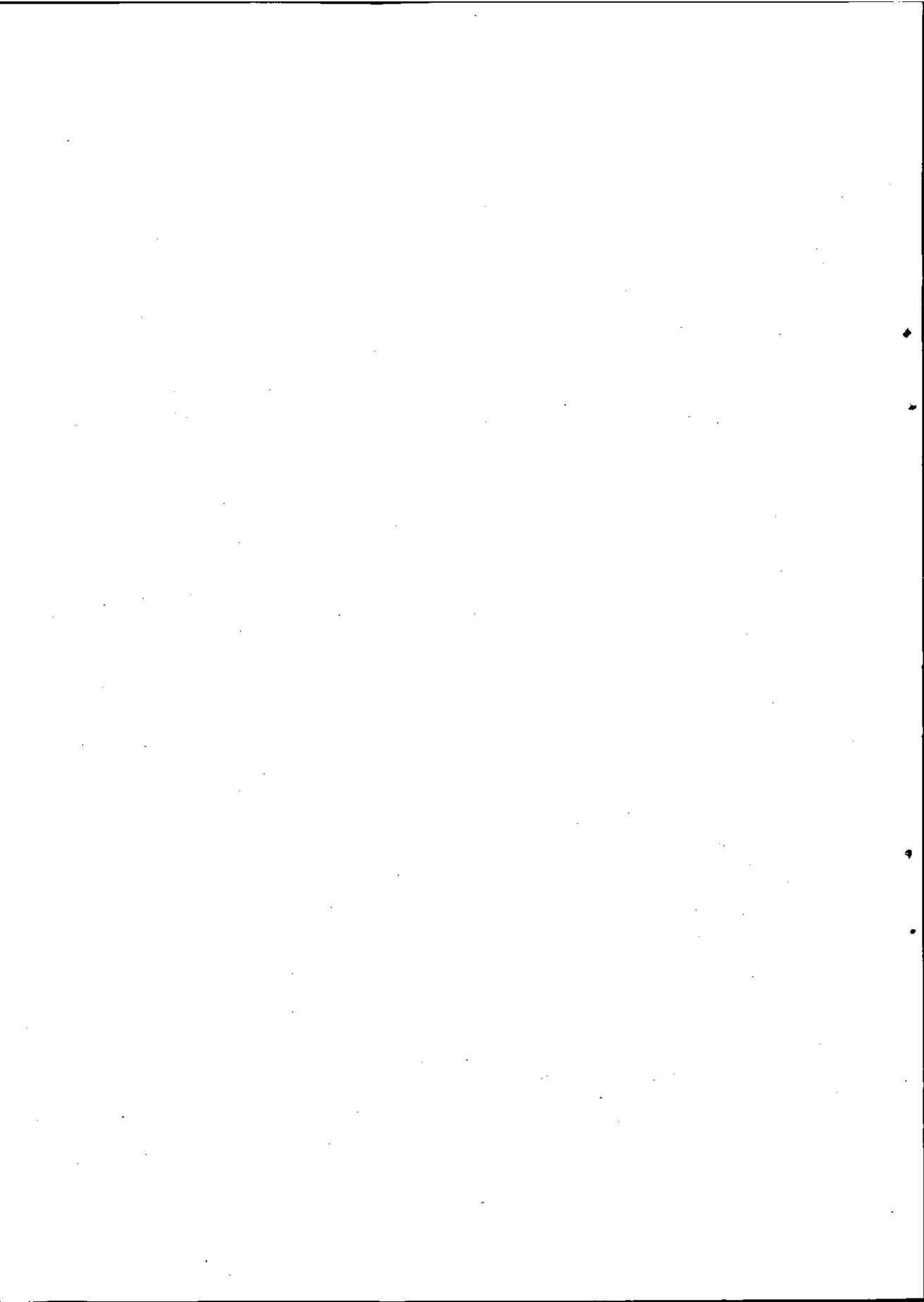
データ・コードの標準化は、鉄鋼情報システムの効率的な運用にとって不可欠の問題である。従って鉄鋼情報システムの詳細設計に取りかかる段階で、データ・コードの標準化の作業を進めるとともに、コード体系の管理体制も明確にする方針である。

###### ③ 蓄積データの利用促進

鉄鋼情報システムの有効利用には、各種の情報の組合わせによる新たな情報の作成、解析の手法およびモデルの開発による情報の質の向上などによる情報の高度利用を促進する開発的な活動も必要である。

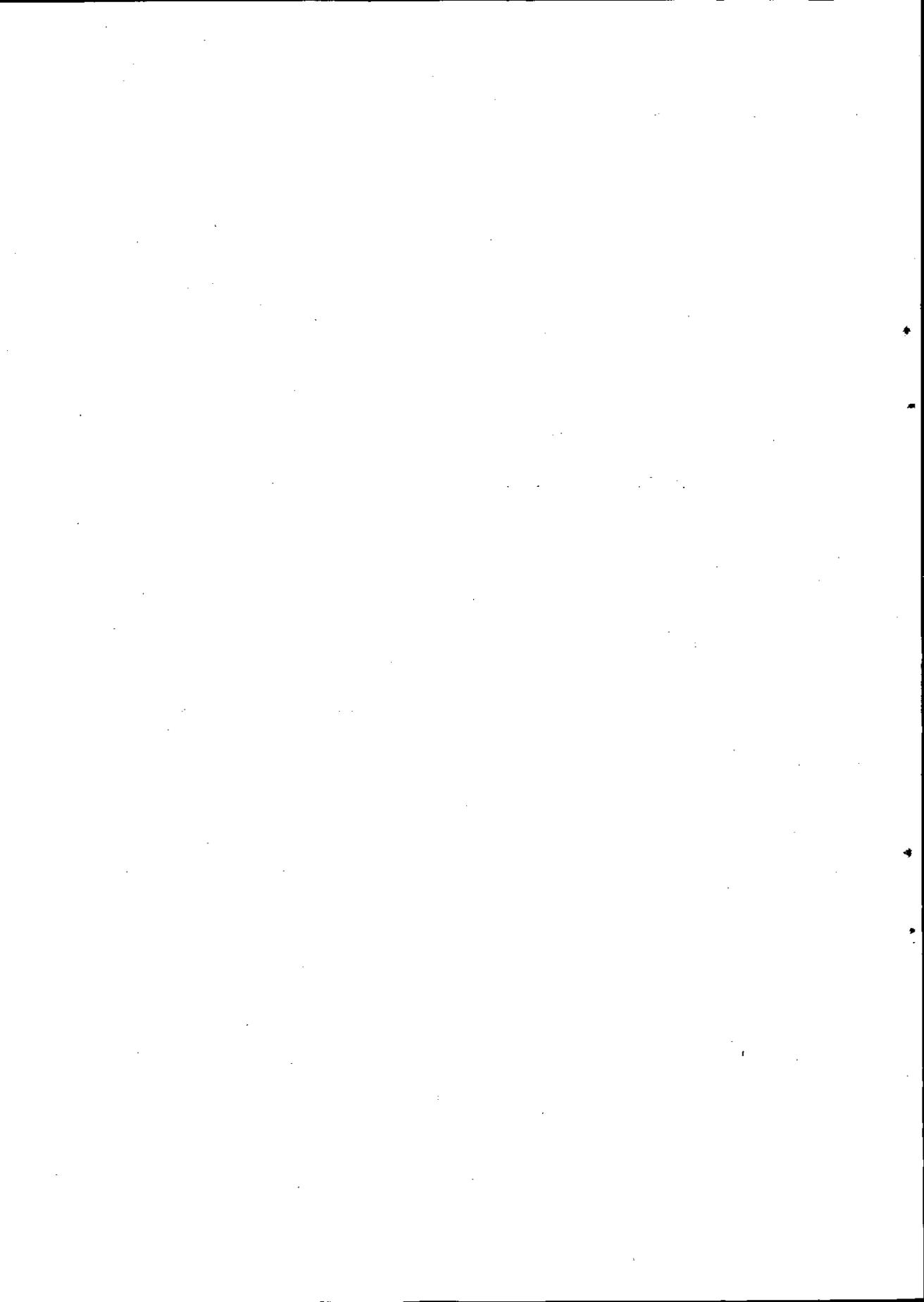
鉄鋼システムのマスター・プラン立案等の企画段階までは情報処理委員会において現状どおり進め、詳細なシステムの細部設計を行なう段階から鉄鋼業界各分野の専門家を加えた構成のS I S運用委員会を設置して進める。

S I S運用委員会には下部組織としてS I Sデータ・コード委員会および、S I S利用促進委員会を設け、それぞれ専門家による検討を行なう体制とする。



#### IV 企業性データバンク

- (A) アソシエテッド・クレジット・サービス
- (B) R・L・ボーク社
- (C) ダン・ブラッドストリート社
- (D) スタンダード・プア・コーポレーション
- (E) マグローヒル社
- (F) 紀伊国屋商店
- (G) 日本の企業性データバンクに関する討論



## (A) アソシエーテッド クレジット サービス

(1) 当社は、銀行、百貨店、商社などを有力株主としており、傘下にCredit Bureau of Greater Houston 及びCredit Bureau Collections を有し、情報提供サービス業はこの両社が行なっている。

(2) Credit Bureau of Greater Houston はヒューストン及びガルフ沿岸地域を対象として、消費者信用に関する情報業務をコンピュータのリアルタイム処理能力を活用して行なっている。

消費者が借入れ、クレジットによる商品購入またはサービス代金の支払いを行なう場合、信用供与者側の信用問合わせ依頼に対して、消費者の支払い態度に関する情報を直ちにできるよう蓄積しておく情報機関となっている。

(3) 信用情報の内容は下記の通り。

(イ) 本人及び配偶者氏名、住所の確認

(ロ) 現職に関する情報

(ハ) 過去の住所録、職歴、配偶者の職業、扶養家族数など

(ニ) 過去のクレジット使用状況、その支払い態度、支配能力と関係する訴訟、判決等の法律関係事項の記録

(4) 信用情報の提供先は、こうした個人信用情報の必要性が明らかな企業でメンバーの資格を取得したものに限定している。メンバー企業は、提供された個人信用情報の取扱いに関する秘密厳守の契約を遵守する義務があると同時に、他の問合わせ先に提供するため自社の有する信用情報を当社に提供しなければならない。このように当社のような信用情報会社に個人信用情報をプールすることによって、信用供与者側は費用を節減できる。と同時に、迅速な信用チェックが可能となり、サービスの向上を消費者に還元できる。

(5) 信用情報社会は、個人信用情報が誤解の余地なく取扱われるための標準用語を以って標準書式で記録・伝達されるよう体制を整備していると同時に、消費者が住所を移動した場合は新住所所管地の情報会社に迅速に移管される体制となっている。

(6) Credit Bureau of Greater Houston のコンピューターは、1969年初からオンライン化が着手され、現在は「第三ゼネレーション・リアルタイム・コンピューター」導入のモデル企業に指定されて「Credipak」という名前でシステム開発が行なわれてきた。

Credipak システムは、IBM360/40、データセルを中核とし、リアルタイムで多数の端末機を有している。

(7) 毎月の処理件数は平均250,000項目の更新の他に新規個人情報平均11,800件の追加が行なわれている。

(8) いわゆるプライバシーの保護に関する問題

信用情報会社2,100社の加盟する業界団体Associated Credit Bureaus Inc. において消費者のプライバシー保護に関する問題を検討した結果、諮問委員会の答申が1968年に提出され、1969年1月正式に採択された。この方針に沿って、それぞれの信用情報会社の手続きが次第に標準化されてきている。現在の信用情報の中には客観的な記録が共通の用語で収録されており、判断は含まれていない。

(B) R. L. ポーク社

当社は創業100年の会社であり、主として都市住所録と自動車関係のデータを有する。

この10年間におけるアメリカの人口は、約3,000万人増加し、1,250万の家屋が建設されており、更には、2,400万台の自動車が新車として全米中を走っているという背景に、当社がコンピューターを活用したデータバンクを形成している存在意義があると考えられる。

現在デトロイト並びにシンシナティにIBM360/65を各1セット設置し主にデトロイトでは統計処理、ダイレクトメール処理を行なっている。何れもバッチ処理である。

以下当社の主要業務と現状における問題点を紹介することとしたい。

a. 都市住所録

当社創立以来の業務であり、個人が誕生するに始まり、結婚、就職、移転、転職、死亡に至るまでの個人データと、企業に関する情報(業種、所有者、経営者等)を収録している。情報収集は5,000人の訪問員が1日1人当たり約100件の先を個別訪問し、約30ページ相当の記入作業を行なって集めている。

データ収集の規模は北アメリカの1,200市、約2,500人をカバーするものであり、本社には支社でパンチされた調査対象者の氏名、住所、電話番号、住居の状況(持家が借家か)、結婚の状況等を内容とした紙テープデータが郵便によって送られてくる。

データは町名、電話番号、或いは職業別に引き出せるようになっており、このデータをもとにダイレクトメールの宛名印刷は勿論のこと、ダイレクトメール実施に当たってのコンサルティングとして、需要層の調査分析、或いは地域別分析等を28人のメンバーで行なっている。

ダイレクトメールの処理方法についても弾力的機械処理が図られており、宛名書き、印刷折り込み、差し込み等のメール作業はコンベア・システム化されている。

また、代理店としてプレミアム券及びクーポン券の商品化或いはギフトの郵送、更には国家試験の請負処理など取扱っているが、単に住所録を発刊するだけではなく、(年鑑として約1,100種の住所録を発刊)顧客—広告—出版という一連の業務を消化しているのが特色といえよう。

#### b. 自動車統計

自動車統計のデータは1946年以降本格的に開始したもので、国家の公式登録資料並びにオーナーリストからデータを収集している。役所関係のデータは1,000項目につき、5~20ドル位の値段の範囲で買っている。

個人の情報と同様に、新車登録に始まり、廃車になるまでの1台毎のアメリカ全国の公式登録関係データ(所有者、車の構造、年次、車型、登録ナンバー等)が収録されており、毎日約3万台のデータ処理を行なっている。デトロイトは周知のとおりフォード、ゼネラルモータース、クライスラーを主体とする自動車工業都市であるが、このデータの主な活用方法はこれ等メーカーの要望に応じ市場情報として、新車の売れ行き、在庫状況等を提供することにある。

また自動車会社からダイレクトメール用のリストを預り、相談に応じてメールも発送している。

なお、情報提供料金は、処理料、コンサルタント料込みのコスト+利益を価格としており、料金は毎年改訂している。

#### c. 銀行関係のデータ

アメリカの銀行関係の財務諸表データ、住所録を保有している。

データの内容は、銀行名、住所、トランジット番号、小切手経路記号、電話番号、設立月日、株価、配当率、経営者、取引先、最近の財務状況等であり、毎年3月、9月には世界の著名な銀行のデータと合わせて銀行名簿を作成している。

この銀行関係の中で特に面白いと思ったのは、銀行ビジネス開発事業部というセクションがあり、ここが銀行と顧客の間に立って銀行に情報を提供していることである。当社ではその情報活動を「対人応答計画」と称しているが、面接者を養成し、銀行の知られていない業務を顧客に説明したり、銀行に対する不平、不満、問題点を卒直に聞き出して銀行に対して毎日報告するというものである。

これなどは、銀行の業務が多様化したことや、或いは著しい機械化に伴って、顧客との情報交換が十分図られなくなってきた皮肉な現象をみた新しい情報処理分野であろう。

なお、現在、当社が直面している問題点は次のようなことである。

(1) 現在当社のデータのCity及びStreet addressには、独自のStreet nameが英文字により72,000分類にコード化されているが、目下これを74,000分類に変更する計画をもっている。

しかし今年(1970年)ビューローオブセンサスが数字による地域コードを付けることに取り組み出しているので、これとの関係が問題になっている。

(2) 都市住所録のデータが年1回しか更新が行なわれていないことである。当社では1回以上行なうとコストが高くなるということであった。

(3) プライバシーの問題に関連し、自動車の登録制度をオクラホマ州が公表しない方向で検討しており、今後のデータ入手に支障が生じるかもしれない。

(4) 個別訪問者の人件費の上昇が著しく、またダイレクトメールを送送する場合の郵便料金の値上りにも頭を痛めている。

### (C) ダン・ブラッドストリート社

当社は年商2億9,200万ドル、従業員22,200人(1969年末)の中堅企業で、マーケティングサービス、出版、企業情報サービスを主たる業務としている。

当社の主要データバンクはマーケティングサービス部門で行なっている「Duns Market Identifiers (略称DMI)」で、約125年前から行なっている社業をベースとして、アメリカ、カナダの約300万件の事業所単位別の企業概要データを収集している。

これは丁度本邦において通産省が収集している「全国工場通覧」に類似したデータであるが、このデータは、経営情報、マーケティングマネジメント情報、セールス情報、マーケットリサーチ、メイリング手段等にその利用目的を置いている。

データの収集は全米で約2,000人のリポーターを常時擁し、又臨時のリポーターを約35,000人動員している大掛りなもので、当社のデータプロセッシングセンターで作成された企業情報ファイルからその一部のデータを抽出してデータ提供サービスを行なっている。

提供方法は、磁気テープ(1レコード350キヤクター)或いは、カードの型で行なわれ、更新データは8週間のサイクルにアップデートされている。

現在の使用電算機は、IBM360/65と2314ディスク4台、磁気テープ20台でダイアリーのバッチ・ジョブを行なっているが、将来はオンラインでマスターファイルを直接アクセス出来るサービス方法を考えている。

なお、当社の提供データの項目並びに業種別収録対象内訳は次のとおりである。

#### I 提供データ項目

名前、ストリートアドレス、メイリングアドレス、市名、国名、Zip コード、  
DUNS ナンバー（当社で付けている会社ナンバー）

1. 地域コードと電話番号
2. 代表者の名前と職位
3. 業 種 名
4. S . I . C . コード第1分類( Standard Industrial Classification  
アメリカ政府の企業グループ分けコード)
5. S . I . C . コード2分類(但し、最高5分類まで)
6. 事業所の従業員数
7. 全従業員数
8. 純 資 産
9. 信用評価基準
10. 販 売 高
11. State - City - Country 地域コード
12. 事業所が本社を示す場合のコード
13. 製造部門が当該事業所にないことを示す場合のコード
14. 事業所が本社所在地のDUNSナンバーをいっしょにもった支社であることを示す場合  
のコード
15. 事業所の親会社のDUNSナンバーをいっしょにもった子会社であることを示す場合の  
コード
16. 細分化した業種、或いは取引体系
17. 設立した年

\* 項目は当社のクレジットサービスを定期的に受けている先におのみ提供。

## II 収録内訳

( 単位千件 )

製 造 業	3 3 5
卸 売 業	3 5 8
小 売 業	1,3 3 4
建 設 業	3 1 2
農業、林業、魚業	2 9
鉱 業	1 8
運輸業、通信業、公益事業	8 7

金融業、保険業、不動産業	34
サービス業	314
計	2,821

#### (D) スタンダード・プア・コーポレーション

米国有数の投資情報サービス社で、企業の財務、株式情報の収集、分析販売を中心に、投資相談、企業調査、定期刊行物の発行を主要業務としている。

1962年からコンピュータ化された財務情報のデータバンク

COMPUSTATは、すでに収録社数3,400社(収録期間最大20年)に達しており、うち1,800社については、FYNALと称するタイム・シェアリング・システムによるサービスが行なわれている。(1969年稼働開始)。また当社が公表する株価指数は、ダウ・ジョーンズ、N.Y.タイムズなどとならぶ米国市場の代表的なものであり、1969年からはこの面でもコンピュータが利用されるようになった。定期刊行物は30種を数えるが、何れも権威のあるものとして高く評価されている。以上の他、最近では、American Bankers Association によって定められたCUSIP System (Committee on Uniform Security Identification Procedure — 有価証券標準化推進委員会の会社名等、記名法標準化システム)のサービス機関としての活動を開始した。

これらのうち、われわれの調査対象はCOMPUSTAT、FYNAL、CUSIPの3者であったが、FYNALについては昨年10月の情報産業議員連盟調査団の報告書に紹介されているので省略する。

なお、当社は1966年、マグローヒル社(McGraw-Hill Inc.)に吸収合併された。またCOMPUSTATならびにCUSIPの業務は当社の子会社Standard Statistics(1962年設立)によって営まれている。

#### (1) COMPUSTAT

COMPUSTATファイルの収録内容は第2表の通りAnnual データとして工業、公益事業、金融業の3種、計3,400社(項目数60~88、年数12~20年)が収録されているほか、工業会社の4年期データ1,800社(項目数17、期数20)が最近追加された。

使用コンピュータおよびMTフォーマットは第3表の通り、CDC-6400、-3000、IBM/360、-7094、-1401計5セット、MTはヘッダー・ラベルつきのものとラベルなしのもの2種で、1部は9トラックも利用可能になっている。

データ・サービスはMTベースで行なわれ、データ更新は、顧客の要求に応じて更新前MTと更新後MTを交換する形で行なわれる。(マスター・テープは毎週更新される)。料金はLibrary FeeとUpdating Feeの2本立てで(900社10年間の場合、それぞれ9,900ドル、4,500ドル—1年)10年以上のデータを必要とするときは追加料金が課せられる。顧客数は全米だけで現在約800とのものであるが、海外については不詳わが国では野村証券と開銀が顧客になっている。

COMPUSTATに収録されているデータは米国SECレポートをもとに再編加工されたものであって、そのためには再分類、加工、チェックの作業が不可欠であるが、この作業は現在も手作業で行なわれており、相当数の労働力(50人)を要する。

Standard & Poors Corp.は前述のようにこのCOMPUSTATをベースにして経営分析、出版コンサルタントなど多面的活動を行なっているのであるが、データバンクとしてのCOMPUSTATそのものは、原データ、それも要約された原データのサービスで必ずしもソフィスティケートされたものでないし、ユーザーの数も全米800程度ではこの種データバンクの顧客数として多いとはいえない。これを利用するねらいは、ややマクロ的に同業種内の比較分析を行なうとか、長期間の傾向分析、あるいは予測ということが中心になるのではないか。

COMPUSTATの採算性を評価することは困難である。しかしこの種のサービス活動が現在全米でほとんどCOMPUSTAT1社に限定されている。収録1社当りの販売価格が基本料11ドルで、更新料1年あたり5ドル程度であることなどからみて、S & P社はCOMPUSTATのサービスを通じて、常識的なFair Return以上のものを得ているとみて差支えないのではあるまいか。

## (2) CUSIP System

CUSIPというのは Committee on Uniform Security Identification Procedure の略で、法人企業の会社名、住所等の記名法を標準化するとともにこれらの名称と社債の利率や満期日等を記した包括的リストを作成、更新して、利用者の便に供しようとするシステムである。American Bankers Association が開発したシステム。

これらのデータ処理には龐大な作業員を要するのでコンピュータが駆使されるが、当 Standard Statistics社は、1968年半ば、このCUSIPのサービス機関として指定された。これまでに45,000社を包含する会社名鑑や、100万以上の市債を記載したリストを発行したが、これらのデータはすべてMTに収録され、毎週更新されて、利用者

の名鑑にも差換え分が送られるようになっている。

CUSIPサービス活動は、まだ始まったばかりであるが、今後はさらに有価証券事務の自動化を目指して、株式ブローカーの一連番号システム、ブローカーの確認、書換え事務の機械化などを計画している。

#### (E) マグローヒル社

世界有数の出版社で、学術書や教科書の出版、ビジネスウィークやケミカルウィークなど約40種の雑誌およびジャーナルの発行、建築関連情報の収集提供サービスなど、広汎な活動を行なっている。企業財務情報の収集提供サービス機関であるStandard & Poors Corpも1966年以来当社の傘下に入っている。

部門別売上および利益(1968)	(売上)	(課税前利益)
	百万ドル	百万ドル
書籍出版	160.9	22.5
雑誌ジャーナル出版	137.2	22.3
情報システム(建築関係情報)	70.0	11.0
S & P社(企業財務情報)		
計	368.1	55.8

従業員数 12,915

事業所数 米国内150ヶ所 海外30ヶ所

これらのうち、S & P社以外では、雑誌の編集と情報システム関係にコンピュータが利用されているが、何れもバッチ処理で、オンライン・システムは未だ採用されていない。

情報システム事業部で発行される定期刊行物は17種に及んでいるが、そのほとんどは建設業関係であり、中でも当社が最も力を入れているのはドッジ・レポートである。

(1) ドッジ・レポートは建設関連業者に対して、入札等の便宜に資するため、41州における建設計画に関する情報を、全米に張りめぐらされたネットワーク(情報収集専門家450人、代理店1100社)を通じて収集(毎日電話または面接によって連絡を保つ)し、これを小型の折たたみ式シートにコンパクトにまとめて毎日提供販売している資料である。調査対象および記載内容は次の通りであって、このための情報処理はコンピュータ化されている。

なお、Dodge Bulletins, Dodge Construction StatisticsおよびDodge Management Control Serviceの3種の刊行物も、上記と同様のデータ・ソースによって作成されている。

## 調査対象

新設または改造で投資額1万ドル以上

業種——商業、工業、協会、住宅、エンジニアリング

## 記載内容

プロジェクトのタイプ名称

建設費概算額

プロジェクトの位置

所有者、建築家、コンサルティングエンジニア、包括請負業者

個別請負業者の名称、プロジェクトの構造、使用資材、機能、内容、能力

その他の重要な情報（たとえば、供託金ポンド所要額、包括請負から省かれた取引条項など）

プロジェクトの現在の状態

以上の他、主な情報サービスには次のようなものがある。

### (2) ブルー・プリント (Dodge/SCAN マイクロフィルム・システム)

マイクロフィルム化された図面を請負業者、建築材料販売業者に提供

### (3) Sweet's Construction & Industrial Catalog System

建設業者のカタログ(4種・26冊)

製造業者のカタログ(3種・計11冊)

インテリア、デザイン業者のカタログ

これらも建設関連業者に販売されている。

### (4) Sweet's Industrial Microfilm Systems

政府民間の設計家、技術者に対する製品や設計に関する情報サービス提供

1967年以降の分で150万ページ(毎年追加分50万ページ)に達するデータバンクを構成し、これらの情報処理によって行なわれている。

### (5) Building Cost Services

建設業者に対するコスト情報として次の4種の刊行物を発行している。

Dow Building Cost Calculator

Dodge Building Cost & Specification Digest

Dodge Construction Pricing & Scheduling Manual

Dodge Estimating Guide for Public Works Const

### (6) Research & Development

コンピューターを駆使して、次に掲げるような各種の技術情報ファイルを蓄積し、蓄積、検索に関する技法を研究中である。

- ① Skidmore, Owings & Merrill 社との提携によって開発された製品技術情報のライブラリー
- ② Sweet's Library System と称する150千ページ以上に及ぶ技術情報ライブラリー、これは建設およびエンジニアリング企業の利用をねらったものである。
- ③ NY州大学と連繋による、建築家が特定の施行基準に合致した建築製品を選定するための情報検索システム

以上のように当社は、多数のユーザーが通信回線の端末を通じて共通のデータファイルに直接アクセスする、という意味でのデータバンクは未だもっていない。多数の人手を要し、必要に迫られた部門から逐次機械化して行く、という通常のコンピュータ化の行き方を堅実にとっているにすぎない。しかし、適用分野がマイクロフィルムとの連動とか、編集印刷機構へのコンピュータ活用、といったかなりユニークなものであることは興味深く、またドッジ・レポートなどは、オンラインではないが多くのインプット端末から毎日入ってくる情報を収集処理し、日報として多数のユーザーにサービスする活動を、効率的かつ迅速に行なっている点からみて、準オンラインともいえるのではなからうか。

## (F) 株式会社紀伊国屋書店

### (1) 企業の沿革

株式会社紀伊国屋書店は、昭和2年創業の総合書店で、東京に本社を置き、都内及び国内主要都市に支店・営業所網を持つと共に、米国サンフランシスコに米国法人の店舗を設立している日本最大の書店の一つである。

国内書籍・雑誌及び外国書籍・雑誌の輸入販売、出版及び事務機、文房具の販売を主業務としており、年商約100億円の企業である。

### (2) 脱活字化社会への対応

同社は、急速に変化しつつある現代の情報化社会への対応として、これまでもマイクロ出版物の販売等に力を注いできたが、ここ一兩年、国内において磁気テープベースの情報への関心が高まりつつあることと並行して、米国を中心にいわゆるMachine-Readable Dataが、どのように流通しているかを鋭意調査してきた。その結果、欧米においては、磁気テープベースのSDIサービス、Retrospective Serchを大規模に行っている団体が加入資格を持つASDICの加盟数が1970年初で34を数えていることからしても、

我国においても企業体が、コマーシャル・ベースで、Machine-Readable Data Baseの販売を行おうとの確信を持つていた。

### (3) Tape, Source

同社においては1971年初、Current Awarenessを中心としたSDIサービスの販売及びMachine-Readable Magnetic Tapeの販売をも併せて行い事業化計画が推し進められ、米国の有力な科学技術情報データ・バンク数社と、日本におけるそれらの独占的乃至は非独占的販売を行き契約を結ぶていた。

主な契約先は、Engineering Index社、CCM Information社（コリア・マクミランの子会社）、American Institute of physics, American Society for Metals等であり、又Institute for Scientific Information社、Chemical Abstracts Service等とも交渉を続けている。

### (4) SDI販売の事業化計画

同社の計画では、1971年7月より営業活動に入る予定で、電算機は国産機HITAC 8210を導入し、ソフトウェアは、そのメーカーのオペレーティング・システムを利用することになる。

事業は3ヶ年を一つの区切りとし、IRの1 Profile当りのコストは2万円～6万円の中になる見込である。

## G 日本の企業性データバンクに関する討論

日本においては、企業性データバンクは、その端緒についたばかりで、いまのところ単にデータ提供者にとどまっているのが実状である。

ここに、日本のデータバンクの萌芽形態であるデータ提供者をわれわれデータバンク研究会が招いて、その実状と問題点及び意見を聴取した記録がある。以下にこれを述べる。

### データ提供者A

#### — 会社概況

特殊性のある科学技術情報をFACOM 230/60で処理している。一応コマーシャル・ベースのSDIサービスつまりCurrent awarenessを配布している。年間3万円という低い価格ではじめたが、この価格では電算機代を出すのがやっとなであり、よく顧客から安くできるなら信頼性がないのではないかといわれる。ともかく1年間続けて来たが、最近では磁気テープで売るケースも2、3件できるようになりSDIとしてはようやく採算ベースに乗ったところである。 )年たない内に大体採算ベースに乗ったが、それなりの苦心があったし、

いろいろとサポートしてくれた専門家の人達がいたお蔭で比較的安くいろいろなことができた。それに年間約7万件しかインプットしていないので数が少ないということもある。また、アメリカ政府の依託研究レポートとか、プロシーディングなどをインプットしているので非常に狭い範囲の特殊なSDIを国内の会員ユーザーに配布している。

#### —インプットの方法

紙テープでインプットしている、オリジナルソースは特になく、他研究機関や図書館の手をわずらわすことなく自分達だけでやっている。雑誌は含まれず2次資料からという訳だが、先に話したのはオービ社の極く一部のSDI業務だけで、その他に会議情報が年間およそ6,000件あり、またグルドルというテープをコンバートして使用している。比率は40%がクリアリング・ハウス、残りの60%がプレプリントとかその他の特殊なものである。

#### —計算機の使用状況

高いが計算センターを使っている。安い値段を付けたのは最初の3~4カ月が試行錯誤の試験期間という意味合いもあった。1年経過した今日そろそろ値上げも考えている。

#### —フランクリン研究所から送ってくるデータの内容

フランクリン研究所はシンク・タンクであるが、われわれは日本と韓国との総代理店ということになっている。同研究所から送ってくる公害情報は年間約2万5,000件である。これは水、大気、土地などに関するものだが、近々騒音とか振動の情報が入ることになっている。これらを合計しても2万7~8,000件であり、一応計算機でインプットしているが、システムが違うので未だ実験段階である。フランクリン研究所では使用しているが、われわれからみると少々雑なような気がする。

SDIシステムは独力で何からなにまで開発した。コスト的に合わなくても業務の1分野にすぎないので勉強のつもりで投資した。

#### —データベースの採算性

誰かが先鞭をつけると、やはり商売だから後からついて行ってもつまらないからやらない。

また政府とか公共企業体のベースは常に遅いし、それに網羅性をもたせようとするから現実的に時間が掛ってしまう。そこでわれわれはその間隙というか、小さな部門を受けもって、しかも危険だから企業の全生命を掛けるようなことはせずに片手間に勉強しておこうというような気持ちでやっている。だから採算ベース以外に何かないと實際上無理だと思う。

例えばU.S.Iシアティック・カンパニーの山川氏などは非常に意欲的に取り組んでいる。ただし人手がないのでどうしてもエージェント的なことしかできないようだ。

データベースは企業の採算ベースにはどうしても合わない。フランクリン研究所のようなこ

とは日本では高くても出来ない。その仕事だけに専門家を100人位抱えている。そういう意味でデータバンクを作ろうとしても難かしいのではない。

われわれのはデータバンクではなく、正直に言えば情報処理サービスを売っているにすぎない。それと営利であるから、サービスをできるだけ早く顧客に届けられるようにフォローアップをしているが、決して自分達のところで全部を持つという方式はとらない。やはりNetworkの拠点になり、海外のものを日本のユーザーに渡し逆に日本のものを海外に流すといったネットワークの一環になるのが精一杯のような感じがする。

#### — 技術情報の問題点

技術情報の場合一番問題になるのは検索手法であり、そしてそれ以前に分析能力である。特許の場合は洩れると困るがコンピュータを使用すれば洩れる可能性は大きくなる。その点苦労があるがわれわれの場合文献だから多少洩れても差支えない。沢山 unnecessaryなものを出せば洩れなくなるが、特許の場合にも洩れを防ぐためには unnecessaryなものを沢山出すことがあるが、経済情報の場合は利用範囲が決まってくるから技術的にもやりやすいのではなからうか。

#### — データ提供における官庁と民間との関係

われわれが商売する場合一番重要なのは速報性である。ところが官庁の場合は機構が大きいため正確ではあろうが速報性が欠ける。このことから民間の方がよいと思う。その代り民間がやる場合はどうしても経済コストを考えるから特異なものを狙うようになる。

官庁がはじめて民間ベースにおろし、民間ベースの速さを維持できれば理想的だがなかなかうまく行かないようである。(JECST)

必ずしも民間が一番よいとは思わないが、たしかにスピードはある。

#### — データバンクの性格

データバンクというのはわれわれのようなものでなく前者をさすのだと思う。しかし、そういうものは営利性が強くないので官公庁でやることになると思う。そうするとスピードのギャップができる。そのギャップをわれわれが埋めるときにわれわれがデータバンク的な性格を一部持つようになるという訳である。

多少のデータバンク的性格といったものは、われわれが国でやるよりも半年ないし1年半早いからその間だけデータバンクであるという訳だ。できれば国に買って貰えれば有難い。

#### — 商売のポイント

結局われわれが商売をする場合に元の情報に何を選ぶかがポイントになる。データ通信になっても同じことであり充分に新企業が起りうると思う。ただ巨大な投資を要すると思うからわれわれはただお手伝いをするということになると思う。

—要望

われわれのようなところに安くコンピュータタイムを提供してくれるようなセンターを設置して貰えないだろうか。つまり公共性のある計算センターで安く実験できるようになれば開発が促進されると思う。補助金を貰ってヒモ付きになるよりも苦しくとも自力で行きたい。

官庁のコンピュータの空き時間、夜間でもよいから貸してくれば、よりよいものができるだろう。

われわれは、日本の情報産業の将来のためにやっているのだから、今後とも育成するように計って貰いたい。

データ提供者 B

—会社概況

われわれのところでは、公共使用と私企業とかMIS使用をイン・ハウスという云い方で分けている。Commercial base へのせようと試みつつあるのをNetwork Information Service という云い方で分けている。米国の例でいうと、上から、Federal Information System (FIS)、それからMISの方をdedicated information systemと呼んでいる。しかし、inhouse という略称の方が世界的にポピュラーだ。次に来るのがNISつまりNetwork Information Systemである。(日本ではNISという、National Information Systemのように誤まって伝えられているが)これは要するにcommercial baseのdata baseである。これをわれわれが毎年評価して日本に売れそうなものを輸入して売っている。

われわれはここ当分の間は計算機を使用せず専らdata baseの輸入元としてやって行くつもりである。われわれのマーケット調査では未だ採算がとれるという結論に達しないのでdata処理をはじめるのは2~3年先のことであろう。

—要望

協会のような指導的立場にある機関にお願いしたいが人の情報を使うということをもっと奨励して貰いたい。とかく日本人は他人を信用しないし、他の情報を使おうとしない。この2~3年こうした風潮は変わりつつあることはある。

それから協関係の点だが、コンピュータ技術以前に、海外情報を利用し、逆に日本のものを海外に進出させようというなら、もう少しそれぞれ処を得しめねばならない。

例えば米国では財務データなど10数年以前からやっていてコマーシャル・ベースにのってからも10年にはなるが、1社しかやっていない。最近になってある会社が財務データに進出しようとしてスタンダード・プア社ははじめて協会会社を持とうとしているが米国では日本の

数倍の経済規模を持ちながら1社でやっている。もともと客観性のあるデータなのに日本では私が知る限りでは12箇所で行っている。

証券業界での用途は全く同じである。外貨の苦しかったとき輸入の試算をしてみたが馬鹿ばかりの数字であり、4社が共同購入すればよいのにと考えた。

#### — 競合関係

われわれの立場からみると興銀の行き方には大変興味がある。アメリカの銀行は商売気が非常に強いが、興銀は財務データバンクを売ったりリディーポルトと組んでいるのとやったりアメリカの銀行と性格が似ている。

アメリカの例でいうと、銀行にはそれぞれ化学オリエントだとかレジャー・ビジネス・オリエントだとか系列があってそれぞれのプライベート・クラブのリーダーシップをとっている。それが一番こわい。

#### — 情報のブランド

日本では情報を自分の名前では売れない。大手の銀行や証券会社は自分の名前では売らざるを得ないが日本ではまだ低い面ではしかとらえられない。

私の社では自分達で開発したもので日本では絶対に自社の名前では売るというポリシーをたてている。外国へ売るときは自社の名前を使っている。

例えば興銀の場合、外注したもので自分の名前を出しているがそれが商売上のネックになっている。

#### — データバンクの成立過程

業務上必要で収集したが情報産業として先鞭をつけておけば儲かるのではないかという思惑があって、いろいろと看板があがっているのではないかと思うが、アメリカの場合は或る程度の補助金を政府から貰って、最初から構想の大きなものと取り組んでいる。

日本の場合には補助金など出ないから、自然会社で出来あがっていたものを少々加工したりむしろそのまま出せるようなものが出ているにすぎないということで、日米間の情報サービスには相当のギャップがある。

しかし地域的な情報を加工したり処理したりする力が元々ない会社が、あらためて情報を買って加工処理して商売するなどということは日本では仲々でき難い環境にあると思う。その点大手の銀行とかメーカーが今までの20億もの投資をしてはじめてわれわれの会社はひとつたりもないと警戒心を持っている。

#### — データニーズ

商売でやっている以上、日米間でパターンの似ているものを探し、その中から採算ベースに

のっているものを選んでユーザー・サイトを徹底的に洗ってから買うことにしている。

やりやすい例ですが鉄鋼などは生産項目が38あり、品目は3種類ある。そこで14数タイプであってしかも共通の生産項目であるから誰でも手に入れることができる。これを使用している工場は米国には30万以上がある。そこで米国ではこのデータベースを扱うだけの市場があるが日本では国がやるより仕方がない。

#### — データサービス

アメリカでは政府もやっているが、民間の方がよいものを出している。ただアメリカの場合は市場があって売れるからやるのであって日本には市場がないだけの話だ。

もちろんコマーシャル・ベースの方がよい。売れなければどんどんつぶれてゆく。8年前からやっているが扱っていたものの半分位つぶれてしまった。小さな会社には自然淘汰が働くが、政府機関から金を貰っているところはつぶれないから自動的な選択が出来ない。

#### — 経営のポイント

必死でやらなければつぶれてしまう。ニーズは作るものだ。とにかく問題意識を探るのがわれわれの仕事だ。その媒体としてコンピュータがたまたまあるにすぎない。

#### — 回線の自由化について

回線の自由化より情報の受け留める傾向の方に私としては関心がある。もしかしたらファミシミリなどがどんどん発達してdata tapeが全然売り物にならなくなるかも知れない。こちらの予測の方をきちんとしておく方が電々公社が何を考えているかより重要だ。巨大なものは国がやってくれるだろう。

#### — 問題点

われわれのレベルが低かったかもしれないが、ともかく日本は人を信用しない風潮があっていけない。

小さな会社であるから1つのメーカーだけと結びついているのは非常に不安だ。メーカーもこわい。

#### データ提供者C

#### — 会社概況

PB社ではオービス社と同じように米国政府のテープを買って、自分のところでプログラムを開発して検索サービスをやっている。売りあげがそう沢山あるとは思わないが、一応これを基礎にしてやっている。

今度アメリカへ行って買って来たものは、技術公社とHEW(保健・教育厚生省)とが共同開発した教育情報のエリック・テープであるが、これは非常に安かった。ですからtape

baseでも売りますし、検索サービスもする。データ数は少なく2万6,000位である。これは米国の大学とか政府機関において電算機を使用しての教育情報はどうかというようなことが入っている。これは渡米前に文部省とか育英会とか東大から要望があった。

現段階では他所から買っているものは2つだけだ。もう1つ現在交渉中であるが、これは32万位のデータが入っている。これを来年買って来て検索サービスを行えば科学技術データは全部揃うと考えている。

その他に、今度は日本の特許の明許書を購入し、内部の専門家が解析し、チェックし、マークシートに転記し、IBMカードにパンチして磁気テープにインプットし、検索サービスを提供しようと考えている。

われわれのところの解析システムを開発し、日本政府におさめたものが2つある。これは今年の2月にスタートしたが受注したものに半導体と電子通信関係の2つがある。そして現在は繊維に取りかかっているが、どうやら今年は採算に乗るだろうと考えている。

設立途上の日本特許情報センターと競合しないようにながりのウェートをしめるような協力をする方向に進めている。

#### — 検索サービスのプログラム開発について

現在プログラムは3つある。I・B・Mと協同で既存のプログラムを1年掛りで改良したものが1つであるが、I・B・Mの場合は基本プログラムを日本でいじれないために、コスト高で時間もかかり、経済的に実施できないということで、今度は国産メーカーを当った。

結局、日立の子会社のNBCが協力してくれて、格安に実験してくれ、プログラムも協同開発している。

#### — データの収集

加工していないデータを集めるといいますけど必ずデフィニションの問題が出て来る。日本の会社の場合はユーザー・アプリケーションだがユーザーの要求が収集する側に反映していないところが現実には最大の問題だ。それをどのようにして解決するかは民間側ではどうすることもできない。やはり収集側がその難かしさを身をもって体験すべきだ。

業務上必要があって収集しているものは既に採算がとれている筈だ。売ればそれだけ儲る訳だから情報産業に進出するということもあるだろう。

#### — 回線の自由化について

回線が自由化されれば結局タイムシェアリングで情報を流す訳ですが、われわれのように特許情報を扱っている場合に、特許屋からみた場合と一般技術者からみた場合とでは観点がかなり違う。特許にたずさわっている者は特許を表現する言葉が3~4あればまず問題がない。

ところがわれわれのような小さな会社の実績をつくるためには安心感を与えるためにできるだけの多くのワード数を使用するようにする。

極端な場合は原文を全部入れるということになる。例えば磁気インクなど使用して出来るだけ質問の仕方をきびしくして言葉の掛け合せを多くすれば自分の探している適格なものがあがってくるという訳だ。

そのようにやった場合、今の電々会社が考えているようなものではタイムシェアリングにならない。キャラクター数で200位にしなければ実際に動かないと思う。

現在は400から600程度でやっているし、化学とか薬学の場合は1つの技術用語でも何百ものものがあるから、もっと沢山入れねばならない。構成要素が沢山あってそれらを掛け合せた場合に何か新製品ができるというようになった場合、今の電々会社の規格ではとてもおこなうことができない。そのような矛盾がある。

秒単位で何故検索するかというと、日本の場合は、特許期間は15年であるが、死んだ特許の再出願を防ぐためには50年位の蓄積が必要である。それにPCTが批准され、サーチ国の数が増えればどこか1ヶ国に出願すればよいことになる。日本だけでも公開制度になれば数年後には企業は相当血眼になって探さねば危険ということになる。だからテープはもちろんディスクでも間に合わなくなる。結局マイクロフィルムとコンピュータが連動するシステムにしなくてはうまく行かなくなるであろうと予測している。

また解析にしても特許にたずさわっていた者と、一般技術者との考え方には大きな差があり、解析をさせると権利をとった人と探す方の人との間にずれが出来てしまう。要するに、すなおに解析しなくてはいけないのだが、特許の弁理士はできるだけ解らないように難かしい言葉を使うので、その解析には大変なお金と時間がかかる。

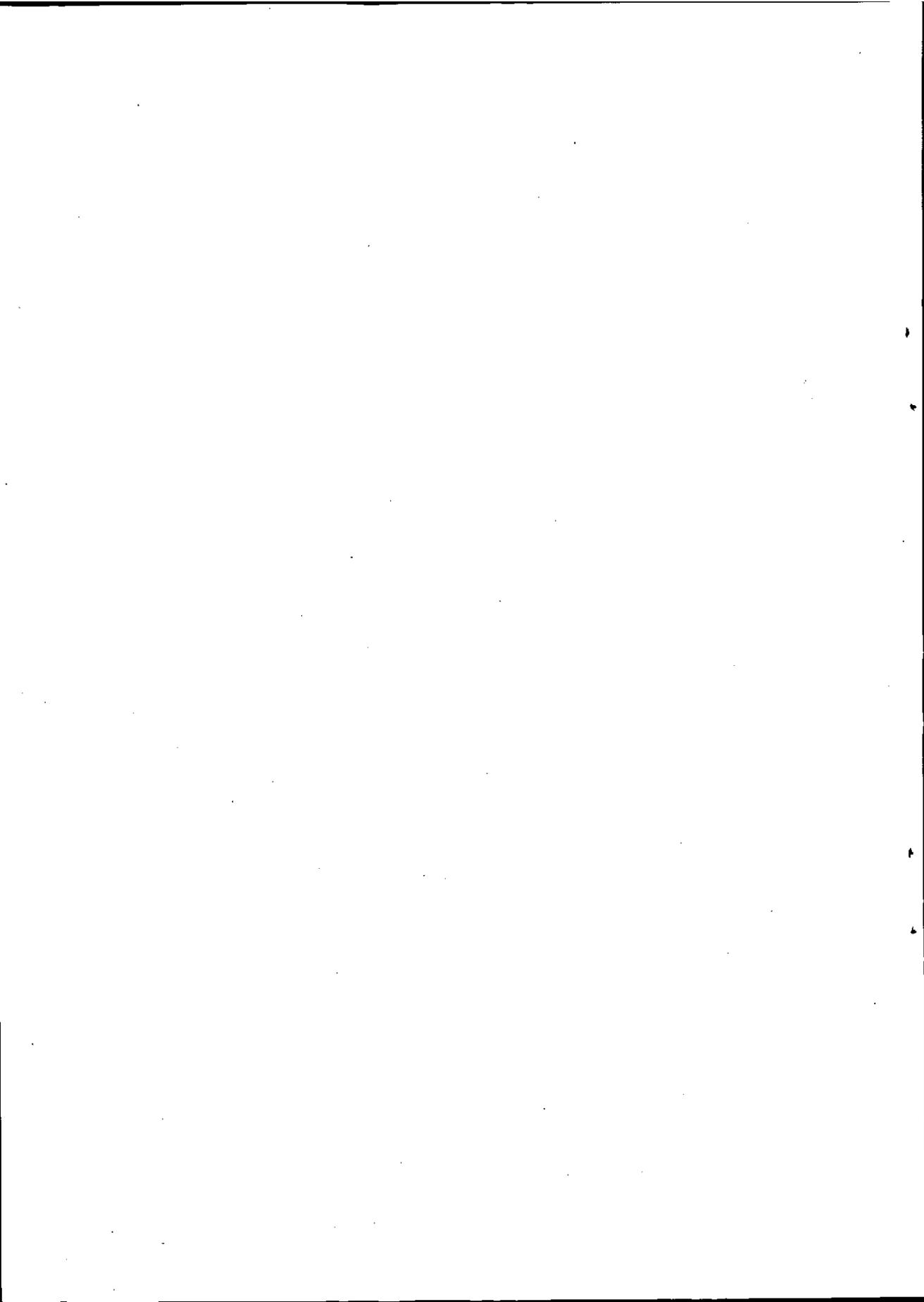
#### — 要望

競争の結果つぶれるのはよいが、アメリカでは政府から8~9割の補助金を貰ってやっているのが多い。日本でもなんとかならないか。私の社ではまだまだ情報を売ってやって行けない。現在やっているのは処理だけである。

## V 公共的データバンク

(A) JETRO

(B) JICST



## 日本貿易振興会 (JETRO)

### 1 転換期における貿易振興策と海外経済情報対策

#### 今後における貿易振興策の基本方針

- (1) 輸出入の大型化、安定化、効率化
- (2) 経済協力の強化、拡充およびその計画性、効率性の確保
- (3) 輸出入相互間及び援助を含めた対外経済活動に関する政策相互間の有機的結合の強化と総合化
- (4) 輸出については、海外需要への一層の適応
- (5) 誘導型、基盤整備型の貿易振興策の推進

などをおかけ、45年度予算要求においても、輸出入、海外投資等企業の海外活動の基盤をさらに整備するため、海外情報収集機能の強化と情報処理体制の整備が第1の重点にあげられている。

#### 今日の先進社会の経済

- (1) 大量消費社会への移行と消費者嗜好の多様化
- (2) 都市開発、宇宙開発、海洋開発などの巨大プロジェクトの出現とこれによる結合市場の出現
- (3) 商品のシステムとしての複雑化、大型化
- (4) 分業化の進展の反面としての分業化された諸活動の総合化の必要性の増大
- (5) 諸活動における国際化の進展
- (6) あらゆる分野の変化の急速化

などに当面し、このような環境の中で、企業は、外部環境に関する適確な、タイムリーな、広い範囲の情報、すなわち、内外の政治、経済、科学技術、市場、他企業の活動等々の情報を社内情報と有機的に組み合わせ、総合的に処理し、これに基づく機動的な経営管理を進めていくことが不可欠となっている。

輸出をはじめとする対外経済活動においても、まさに情報化の要請は、そのすべての前提になるものであり、海外経済情報の充実、強化は、かねがね叫ばれてきたものである。

### 2 海外経済情報需要の動向に情報ネットワークの必要性

海外情報が、特に国際取引の推進にあたってぜひとも必要なわけは、国際取引自身の特色に由来している。

#### 国際取引の特色

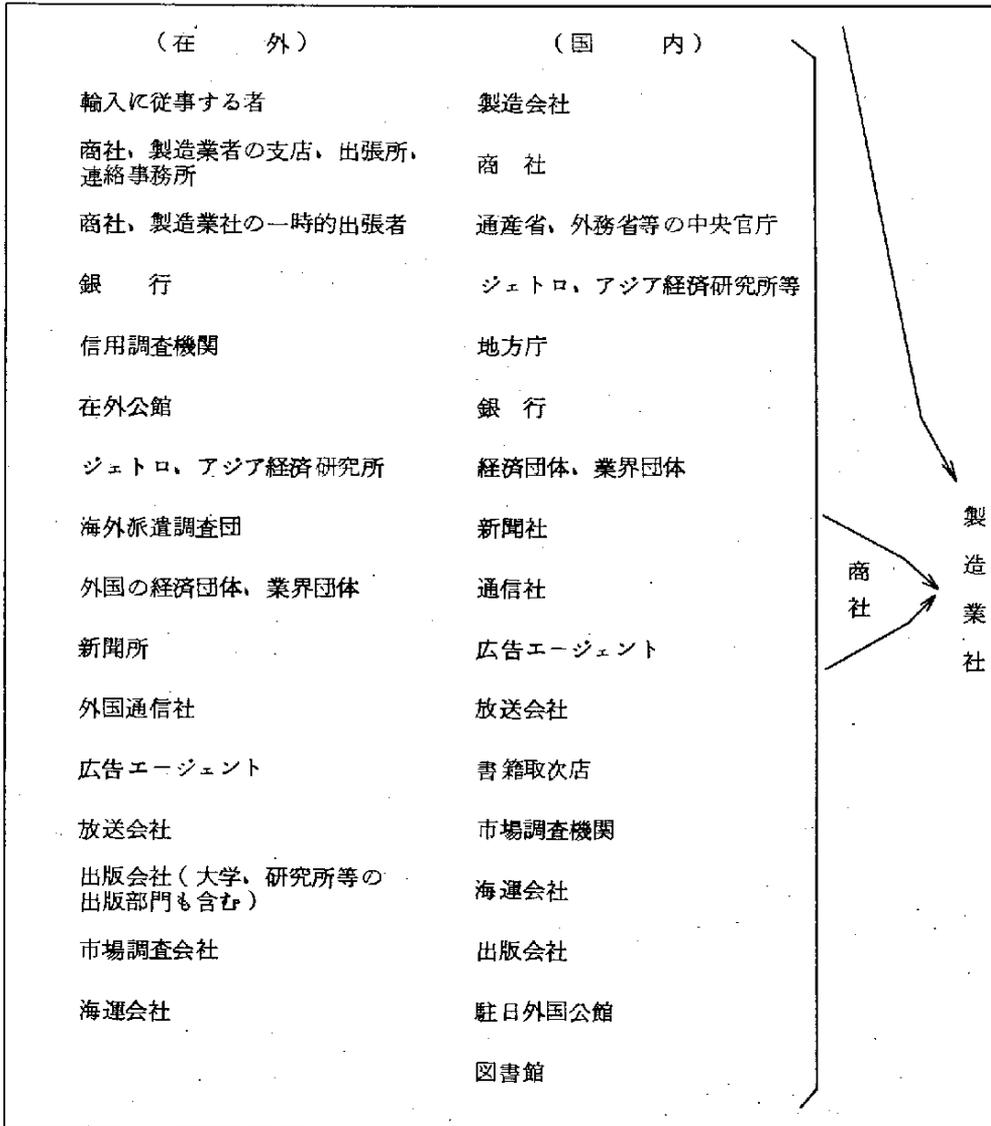
- (イ) 地理的に遠隔地間の取引であること
- (ロ) 商慣習、取引制度の異なるもの間の取引であること
- (ハ) 趣味、嗜好の異なる市場を対象とする取引であること
- (ニ) 支払手段が異なること
- (ホ) 言葉が異なること

ジェトロをはじめとする海外経済情報機関は、まさに

- (イ) 取引成立に直接必要なもの
- (ロ) 取引にあたって入手しておいた方がよいもの
- (ハ) 入手していると取引上有利なもの
- (ニ) 取引の機密性から公開情報だけに頼っておられないもの

の拡充をねらいとするものであり、海外情報収集機関のネットワーク確立が望まれる所似でもある。

V-①表 内外における海外経済情報収集源



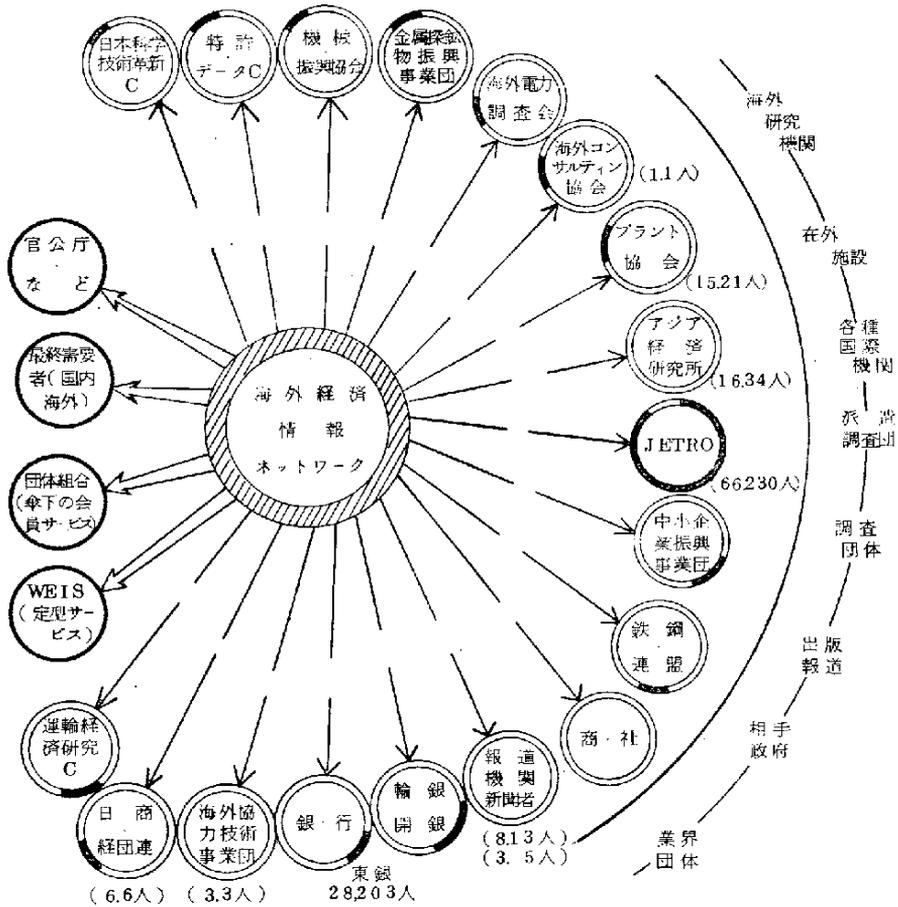
第1表には、メーカーや商社がどれほど多様な情報源をもっているかを示すために、在外および国内における情報源をアトランダムに拾いあげたものである。

V-1図は、海外経済情報を収集、提供する種々の機関を集めたものである。

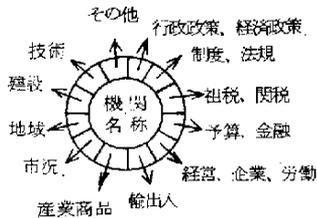
海外経済情報ネットワークという、海外経済情報システムを確立するということは、簡単にいえば、この図に現われてきた機関を如何に有機的に結び、ほんとうに必要な情報が需要

### 3. 海外経済情報機関と海外経済情報ネットワーク

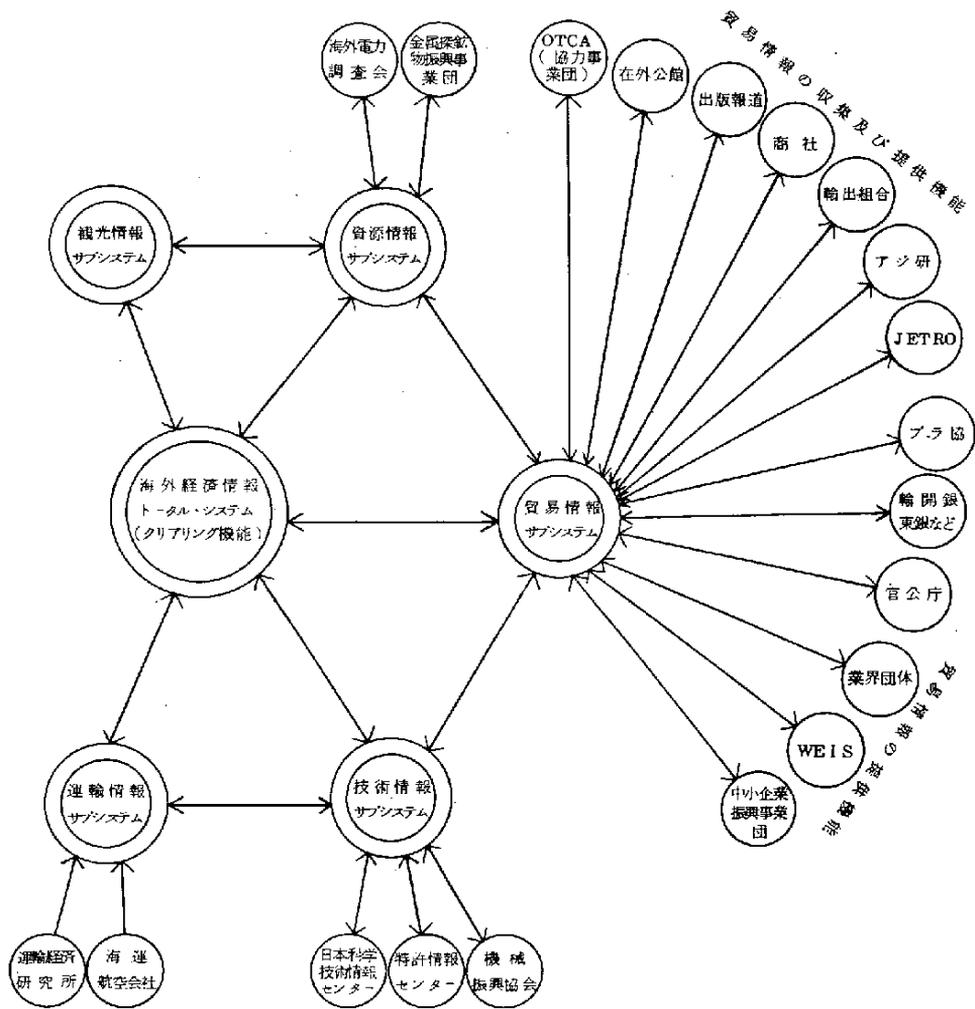
V-1 図 海外経済情報の収集 提供ネットワークと情報内容



- (凡例)
- 1 機関名外側の ( ) は当機関の海外施設及び派遣人員を示す。44年8月現在
  - 2 枠内の黒字は情報のガバレッジを示す。ガバレッジ区分は左図のとおり。



V-2図 海外経済情報システム図



者に迅速に、適確に提供されるようなシステムをつくりあげることである。

したがって、海外経済情報システムが効率的に運用されるためには、いくつかのサブシステムに分けてそれぞれのサブシステムの完成からはじめなければならないだろう。

V-2図は、予想される海外情報システムの全体像を描いたものである。

トータルシステムとして海外経済情報システムは、貿易情報サブシステム、資源情報サブシステム、技術情報サブシステム等々により構成されている。

#### 4 通商産業行政をめぐる情報ネットワークと海外経済情報システム

##### 施 策

- (1) 情報処理の普及、高度化のための環境作り（電子計算機利用の普及、高度化）
- (2) 所管産業の情報化、システム化の推進
- (3) 通商産業行政をめぐる情報処理の普及、高度化と産業界等に対する情報提供

次に示す第三図は、通商産業行政をめぐる情報ネットワークの構想とこの中における海外経済情報システムの位置づけである。

## (B) J I C S T

### 1 沿革

J I C S Tは昭和32年8月、科学技術庁の最初の政策の1つとして、それ自身を造るプロジェクトによって設立された。

最初、役職員62名、資本金80000千円で出発し、昭和33年よりサービスを開始し、昭和46年度において、次の事業規模をもつ。

V-②表 収入支出総括表（単位：千円）

収入	科 目	46年度認可
	1. 国庫支出金	1,038,000
	2. 事業収入	886,986
	1) 出版収入	483,436
	2) 受託サービス収入	397,650
	3) 講習会等収入	5,900
	3. 事業外収入	23,990
	4. 前年度繰越金	6,060
	収 計	1,955,036

支出	科 目	46年度認可
	1. 収 集 整 理 費	408,779
	2. 提 供 事 業 費	438,002
	3. 一 般 管 理 運 営 費	915,775
	1) 人 当 経 費	535,562
	2) 特 別 経 費	125,601
	3) 業 務 合 理 化 促 進 費	254,612
	4. 管 理 施 設 費	127,480
	5. 予 備 費	65,000
	支 出 合 計	1,950,377

V-③表 収集処理データ量(年令)

摘 要	46年度認可	備 考
収 集 外 国 雑 誌 数	5,200種	環境公害文献100種を含む。
収 集 国 内 雑 誌 数	2,300種	
収 集 外 国 特 許 数	4,800件	化学(米、英、独)38000件、電気(米)10000件
収 集 医 学 雑 誌 数	0種	
収 集 二 次 資 料 数	13種	
収 集 レ ポ ー ト 数	7,500件	原子力レポート7,500件
処 理 外 国 論 文 数	28,200件	原子力レポート3,000件を含む。
処 理 国 内 論 文 数	7,300件	
処 理 外 国 特 許 数	3,800件	化学のみ
処 理 中 小 企 業 海 外 技 術 情 報	600件	
処 理 食 品 工 業 技 術 情 報	600件	
処 理 海 外 技 術 ハイ ラ イ ト	2,400件	
処 理 環 境 公 害 文 献 集	4,040件	理工学文献2000件、医学文献2000件、トピック情報40件

これは、前身となる機関も、設立推進母体もなく、全く新しく、科学技術振興政策によって造られた機関である点に大きい特徴がある。

法人格は、日本科学技術センター法により設立が定められた。科学技術庁の監督下にあるいわゆる特殊法人である。

## 2. 業務の内容

(1) 国内外の科学技術文献を収集し、重要な論文記事を抽出し、抄録、キーワード索引を作成する。

(2) それを漢字穿孔鍵盤を通して入力し、コンピュータにより編集し、高速漢字出力装置によって抄録雑誌の版下を作成し、印刷する。

その分冊は次の通りである。

### 科学技術文献速報

- 工学一般・機械工学編 (半月刊) [昭和33年3月創刊]
- 電気工学編 (半月刊) [昭和33年3月創刊]
- 化学・化学工業編 (旬刊33~39年度半月刊) [昭和33年3月創刊]
- 金属工学・鉱山工学・地球の科学編 (半月刊33.34年度月刊) [昭和33年9月創刊]
- 土木・建築工学編 (半月刊33.34年度月刊) [昭和33年9月創刊]
- 物理・応用物理編 (半月刊34年度のみ月刊) [昭和34年4月創刊]
- 原子力(アイントープ)放射線利用編 (月刊) [昭和36年4月創刊]
- 経営管理編 (月刊) [昭和38年4月創刊]
- 国内化学編(日本化学総覧) (月刊) [昭和39年1月創刊]

### (3) その他の出版物の刊行

(i) 海外技術ハイライト (月刊) [昭和42年4月創刊]

一般的な展望記事、新製品の紹介、各国の科学技術の動向、統計資料等、珍しい興味あるニュースを掲載したものである。

(ii) 外国特許速報 (週刊) [昭和43年9月創刊]

外国特許速報は、米国・英国・西独3ヶ国の化学工業に関する特許明細書を網羅的に収録し、和文の標題によって簡潔に内容を示したものである。昭和33年4月創刊、週刊で刊行している。

化学部門に関するかぎり、これは米国のオフィシャル・ガゼット、英国のオフィシャル・ジャーナル、西ドイツのペテント・プラットフォームと同等のものである。

(iii) その他

情報管理、中小企業海外技術情報、食品工業技術情報、日本特許索引、情報カード、特殊テーマについてまとめた文献集等を発行している。

(4) 依頼に応じて行なうサービス

出版物により提供された、あるいは任意に集められた技術情報は、利用者にとり取捨選択されてさらに詳細な資料の要求となつてはねかえてくる。この要望に応えるものが受託業務であり、JICSTでは複写、翻訳、調査、抄録等のサービスをおこなっている。

(a) 複写

文献複写は、JICSTの所蔵する資料の複写、外部機関および外国機関の所蔵する資料の複写、文献の目次の複写等をおこなっている。

(b) 翻訳

重要必読の情報の社内流通伝達のために、また特殊外国語の読解や海外への紹介等のために、学術論文、特許明細書、カタログ、通信文その他の翻訳をおこなっている。なお外部の翻訳協力者は2567人に及んでいる。(6月2日現在)

(c) 調査及び抄録

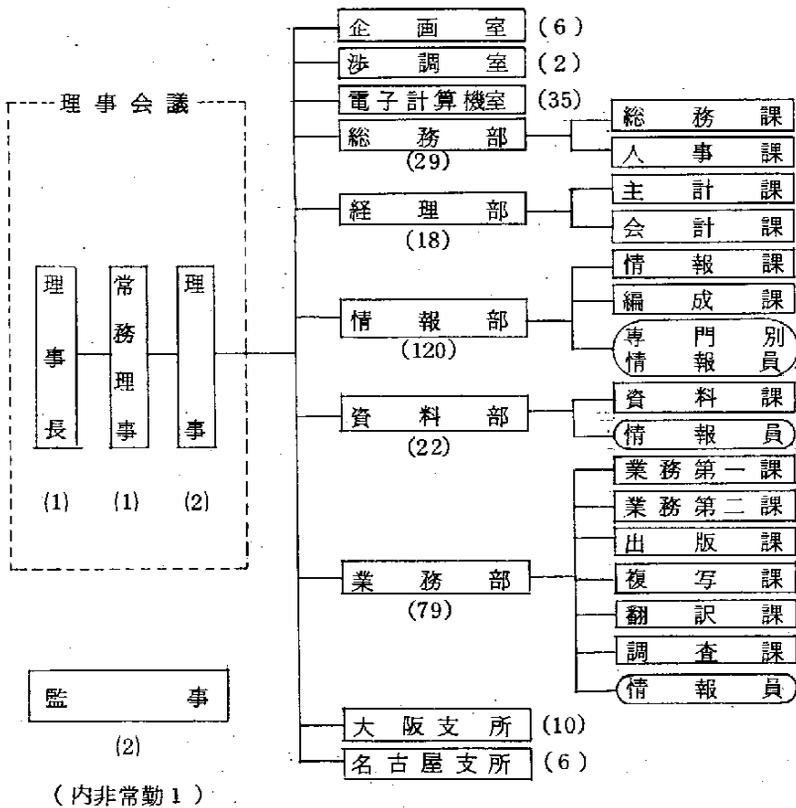
医学、農学、生物学を除く理工学全分野の特定の主題に応じて、過去の情報を検索調査し、文献および特許のリストの作成、入手困難な文献の取り寄せ、抄録作成等のサービスをおこなっている。

3. 組織および保有施設

(1) 組織

日本科学技術情報センター機構図

( )の中は昭和45年度認可定員



(2) ハードウェア・システム

(i) 中央処理装置

FACOM 230-50

主記憶容量

65 KW

補助記憶および

磁気テープ装置 8台

入出力装置

磁気ドラム装置 1台

磁気ディスク装置 2台

カード・リーダー 1台

ライン・プリンター 1台

(ii) 中央処理装置 FACOM 230-25

主記憶容量 45KB

補助記憶および 磁気テープ装置

入出力装置 磁気ドラム装置

磁気ディスク装置

カード・リーダー

ライン・プリンター

(iii) 高速漢字出力表示装置 GEM 3800

漢字の字種およびプリンタの性能を下表に示す。

(V-④表)

字 種	インプット	アウトプット
漢 字	1861	1861 (明 朝)
カ タ カ ナ	81	162 (明朝、ゴシック)
ひ ら が な	77	154 (      )
英 欧 文 字	65	195 (ローマン、イタリック、ボールド)
ロ シ ア 文 字	66	198 (      )
ギ リ シ ヤ 文 字	33	99 (      )
ア ラ ビ ア 数 字	10	30 (      )
ロ ー マ 数 字	20	40 (ローマン、イタリック)
記 号	199	248 (ローマン、一部イタリック、ボールド)
ス ペ ー ス	6	6
予 備	78	78
計	2496	3071

字 種	フィルムレコーダ	ソフトプリンタ
記 録 媒 体	フィルムまたは印画紙	銀塩安定化印画紙
媒 体 巾	9.2 mm	182 (B5)、210 (A4) mm
媒 体 長 さ	1 0 0 m	1 0 0 m
プリント文字大きさ	6 7.5 10.5 ポイント	10.5 ポイント
1 行 の 長 さ	最大 30 インチ (126 mm)	最大 30 インチ (126 mm)
行 間 隔	1 ポイント単位で可変	1 ポイント単位で可変
印 字 速 度	5~10 行/秒 (200~600 字/秒)	5~10 行 (200~350 字/秒)
現 像	外 部	湿式内蔵
解 像 力	20 本/mm 以上	

### 3. ソフト・ウェア・システム

#### (i) JICST 総合 IR システム

- (a) 漢字モード
- (b) BCD/EBCDIC モード

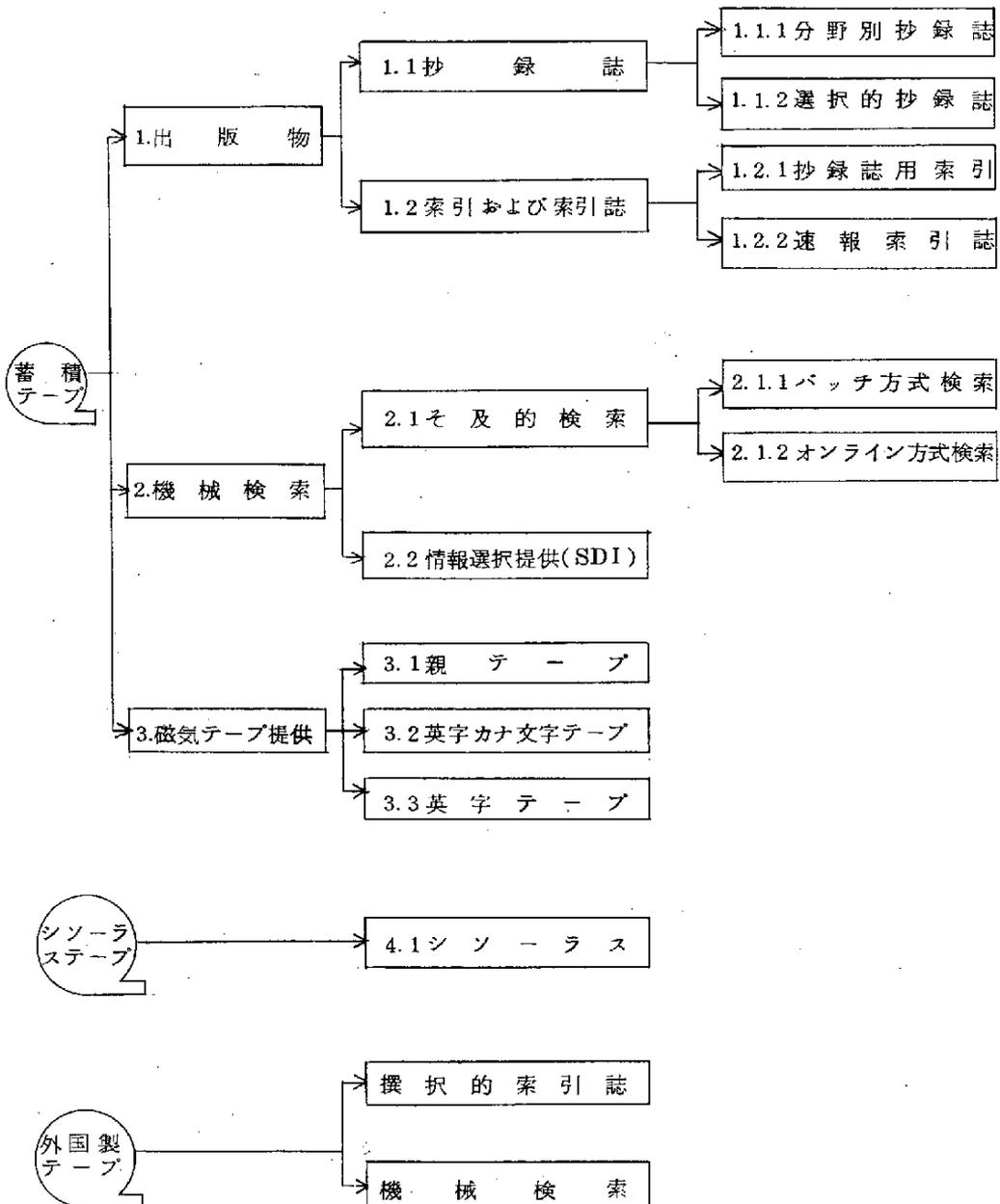
#### (ii) JICST 用 管理システム (DOCTOR)

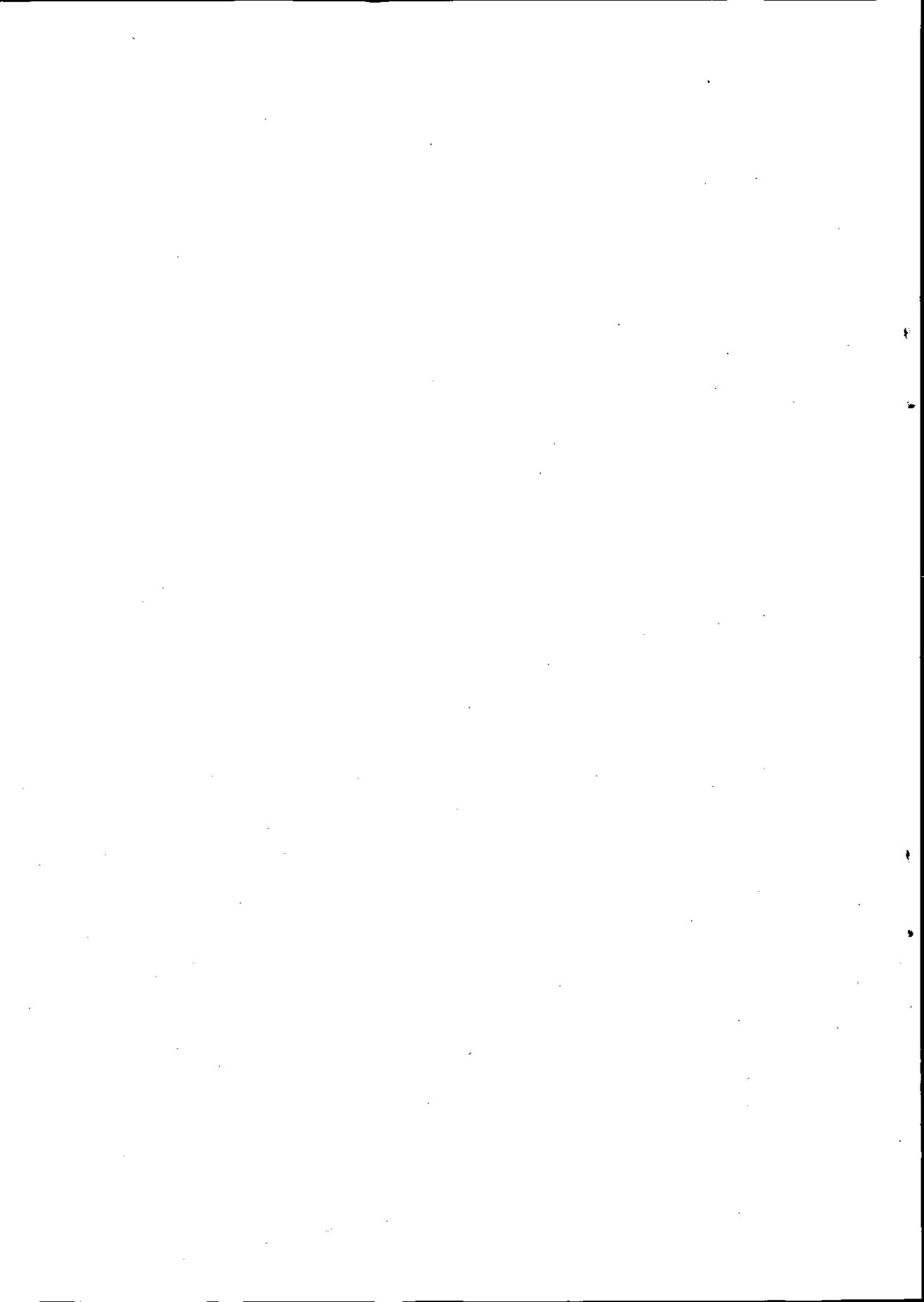
- (a) 漢字モード
- (b) BCD モード

#### (iii) 科学技術文献速記編集システム

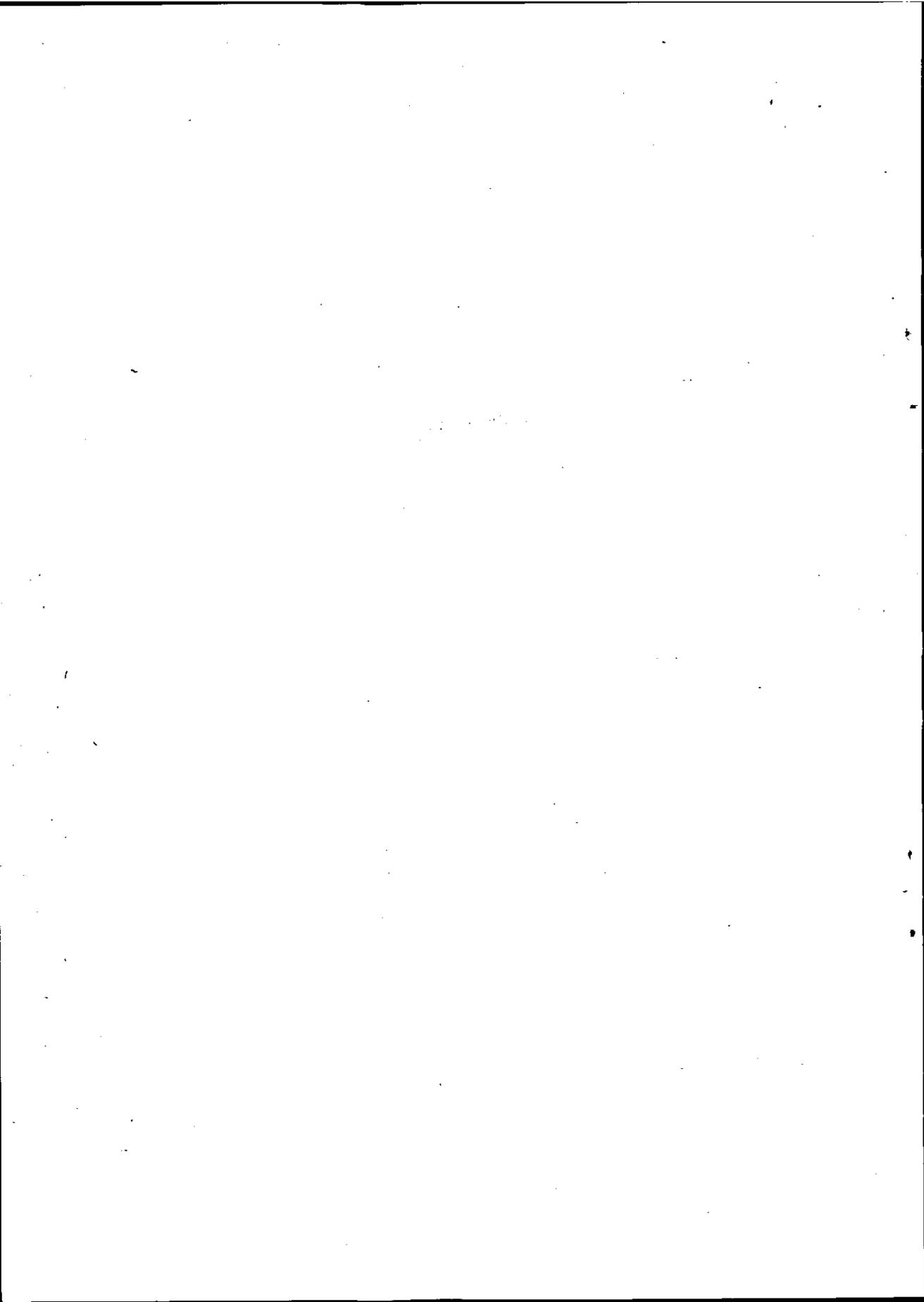
V-5 図 これらのソフト・ウェアによるサービスの概念図

V-5図 これらのソフト・ウェアによるサービスの概念図





## 第3部 データバンクの分析と評価



われわれは第1部でデータベースの構造を理論的、技術的に明らかにした。第2部でデータベースの現象を数多くの事例について調査した。

ここでは、データベースにつき理論的研究と事例研究を媒介として現状分析を行い今後のあり方について評価を加えた。

まず第1に、コンピュータの発展がデータベースの成立にどのような影響をもたらしたかを検討する。

第2に、機能面からデータベースを評価する。

第3に、データベースはどのような生成発展の過程をたどって成長してきたかを考察する。

第4に、日本、西独、米国のカリフォルニア州のパブリックセクターにおけるデータベースの現状比較を行う。

われわれは、以上の視点のほかにもいくつかの問題、たとえばデータベースの経営分析、情報価格形成論など実証的研究を行ったが、データ不足など色々な事情があって、成果を挙げる事ができなかった。これらの問題は今後の研究を待つことになる。

## I コンピュータの発展とデータベース

本節では、コンピュータ利用の相違がデータベースの種類、利用の仕方、経営などに大きな影響を与えるという認識から、コンピュータ利用発展とデータベースとの時間的相互関係を考察する。

### (1) コンピュータ利用の発展

図31はコンピュータの導入が歴史的経過にしたがって、日本とアメリカとで、どのような相違が生じてきているかを示している。

コンピュータ利用の発展段階をわれわれは図3.1に示した如く、次の6段階に分けた。

第1段階MS：マニュアルシステムを利用する段階

第2段階PCS：パンチカードシステム(PCS)を利用する段階

第3段階EDPS：エレクトロニクスデータプロセッシングシステム(EDPS)を導入する段階

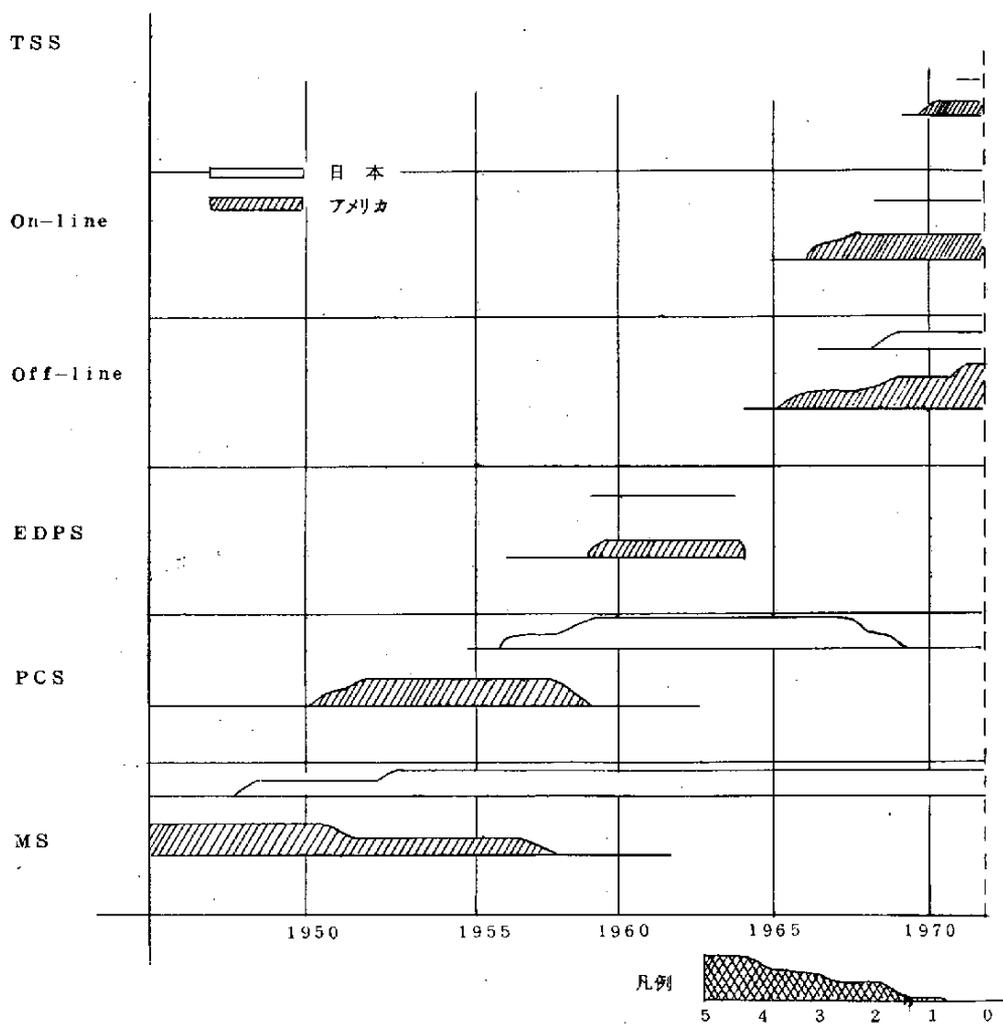
第4段階off-time：ランダムアクセスメモリとオペレーティングシステムを利用する段階

第5段階on-time：コンピュータとデータ回線との結合をはかる段階

第6段階TSS：タイムシェアリングシステム(TSS)の普及する段階

第1段階と第2段階はコンピュータ以前のデータ処理システムであるがこの段階においても第1部で述べた定義におけるデータバンクが成立するのであえてひとつの段階として分類した。

図3.1 コンピュータ利用の発展段階



そこで図 3.1 にしたがってコンピュータ利用の発展段階における日米の相違を歴史的に跡づけてみると、EDPSの開発時点は、アメリカでは1956年より始まり、日本は1959年より開始した。その差は3年である。オペレーティングシステム(ランダムアクセスメモリー)の利用開始期はアメリカが1963年、日本が1966年で、その差はやはり3年であった。たしかにオペレーティングシステムの導入の時期については日本とアメリカとに大きな差はなかったと我々の経験にてらして言えるが、しかしその量的展開は日米経済規模の大小で標準化したとしても、日本と比べてアメリカは格段な大きさであった。

さらにまたオンライン化やTSS(タイムシェアリングシステム)の利用にいたっては、日本ではデータ回線の自由化が遅れていることもあって、この面では相対的立遅れを呈している。

では日本が相対的に遅れているTSSのアメリカにおける利用はどのような状況であろうか。表3.1は、アメリカにおけるTSSの利用状況をあらわしている。

表3.1 アメリカにおけるTSS利用状況

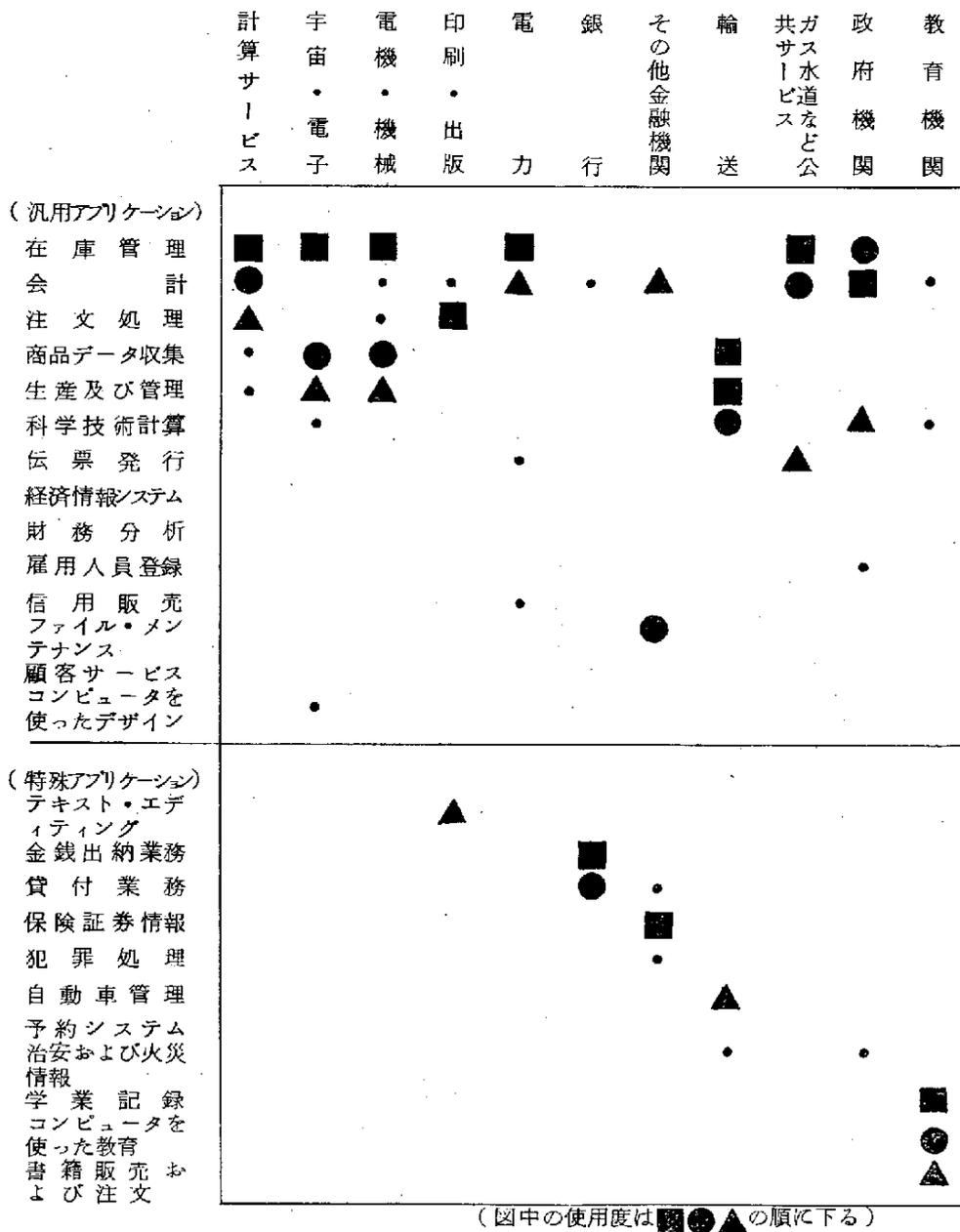
業種	宇宙	銀行	通信	化学	石油	電子	コンサルティング	教育	政府	輸送	製造工業	公益事業	サービス業	医療	金融	鉱業	建設	食品、飲料	出版	その他
平均費用	113	108	89	82	55	55	54	53	52	41	26	26	23	22	19	10	10	4	3	79
最大利用のユーザの(社)	7	8	6	3	4	7	4	5	10	3	5	3	4	3	4	3	2	1	2	3
社平均利用数(社)	1.7	2.3	2.7	1.8	2.2	4.0	1.7	3.1	3.2	1.7	1.9	1.7	1.8	1.8	2.2	3.0	1.8	1.0	1.5	2.3

表3.1をみても明らかなように、アメリカでは沢山の数ではないけれどほとんどの業種において何らかの形でTSSを利用しているが、図3.2に示されているように、その用途は在庫管理などのごく一部にいまのところ限られている。しかし、こうしたコンピュータ利用の

相違がひいては情報利用の日米格差の原因ともなっており、具体的にはデータバンクの種類・利用の仕方、経営などの面に日米ギャップを生ぜしめている。

また、図3.2は、アメリカにおけるTSS利用の用途およびその範囲を示している。

図3.2 アメリカにおけるTSS利用の用途とその範囲



(2) ハードウェアの開発とその普及

ハードウェアの開発経過は図 3.3 に示されている。

図 3.3 ハードウェアの開発

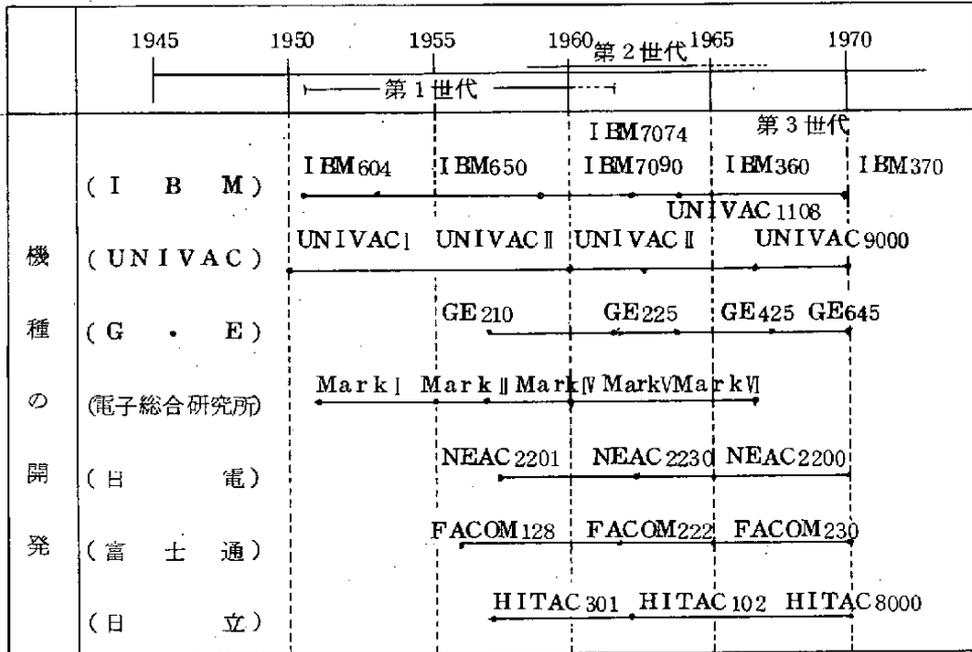


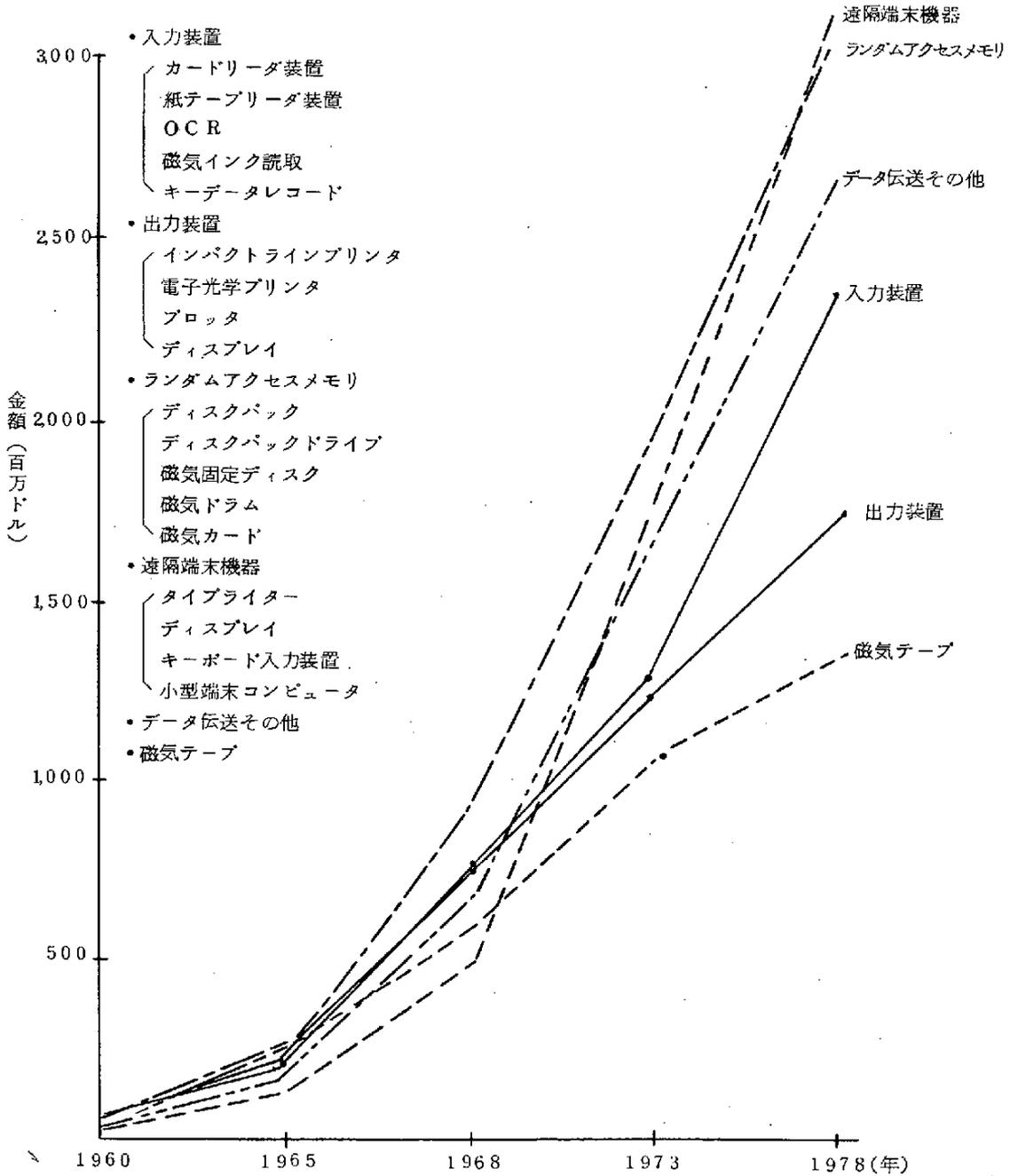
図 3.3 にも明らかなように、第 1 世代のコンピュータの開発期には、アメリカでは、IBM 604 を、それにきびすを接して日本では Mark I (電子総合研究所) を開発していた。また 1950 年代の後半から 1964 年までの第 2 世代のコンピュータの展開期には、IBM 650 や UNIVAC II に対抗して、日本では NEAC 2201、FACOM 128、HITAC 301 などの国産コンピュータは開発されてきた。

IBM 360 が発表されて第 3 世代のコンピュータ時代に突入したのちも、NEAC 2200、FACOM 230、HITAC 8000 と数年を待たずに日本でも第 3 世代のコンピュータを開発してきた。

これらの事実から、ハードウェアの開発時期及びその速度は、日本とアメリカとはさしたる差はなかったといえよう。

図 3.4 は各種機器ならびにデータ伝送のアメリカにおける現在の普及状況と将来予測を示したものである。

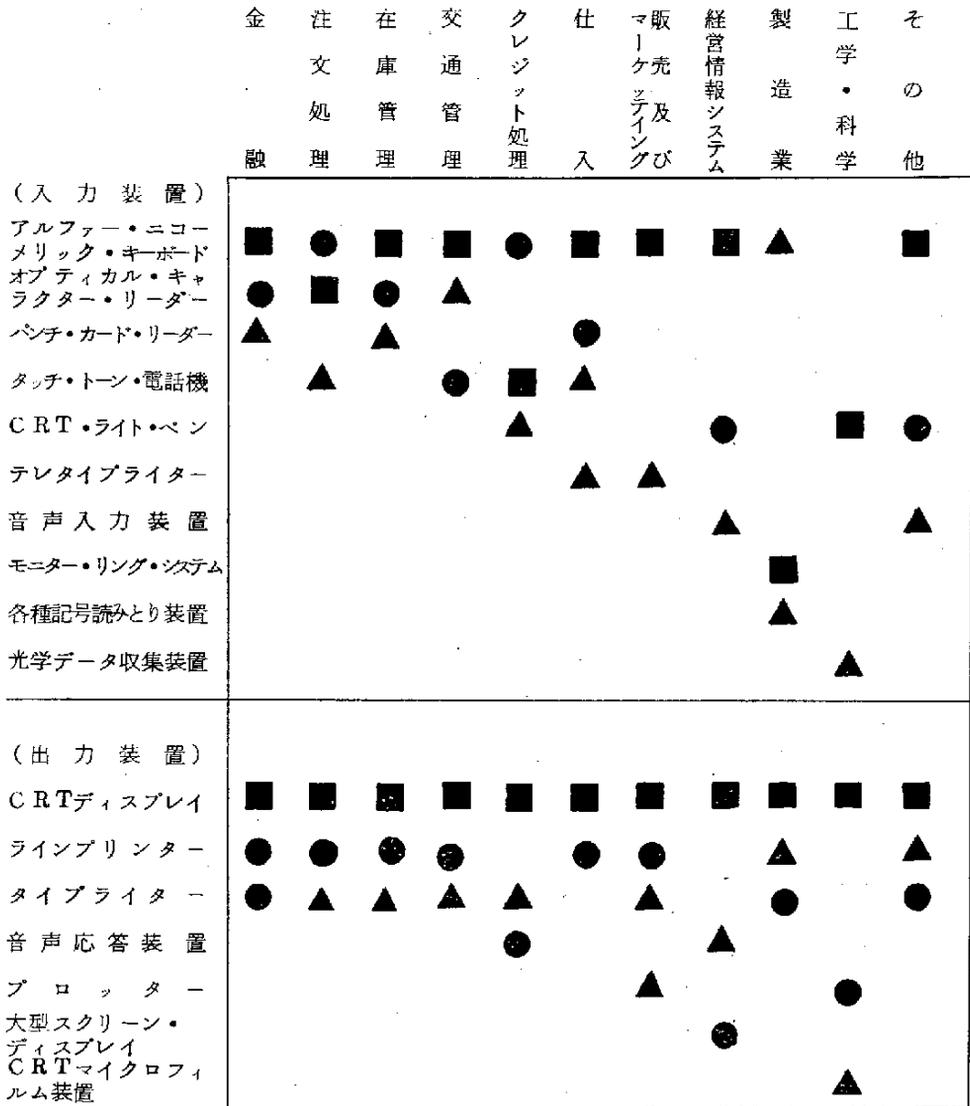
図 3.4 アメリカにおける周辺機器種類 — 普及の推移



(出所) 機械振興協会経済研究所  
米国の電子工業需要予測

アメリカにおいて、インプット・アウトプット面での端末機の普及が日本とくらべ著しいわけであるが、その用途及び範囲に関する実態が図3.5 アメリカにおけるターミナル利用の実態に示されている。

図3.5 アメリカにおけるターミナル利用の実態



(図中の使用度は■ ● ▲ の順に下る。)

図 3.5 によれば、インプット・ターミナルではアルファ・ニューメリック・キーボードの利用が著しく、アウトプット装置では、CRTディスプレイの使用度が大きい。

### (3) ソフトウェアの発展

データバンクの成長を促進する要因として前述したEDPSの普及、およびハードウェア、ソフトウェアの発展があげられる。とりわけデータ・マネージメント用ソフトウェアの開発が非常に重要である。そこで、アメリカのデータ・マネージメントシステムの現状についていささか述べてみることにしよう。

ここでいうデータ・マネージメントシステムとは、総合データ・ファイルの作成、更新、修正等のデータ管理と、ユーザーのアクセスに対する秘密保護、検索、及び検索データの分析、作表等の機能を総称したものであるが、この芽生えは、1956年にIBM702用に作成されたMARK I Generaliyad Report Generator にみられる。一方軍の情報システム用としてTUFF-TUG-IRSがIBM704用に開発され、これがFFS、NIPSと発展し、最近のGISにつながっている。現在多くのデータ・マネージメント用ソフトウェアが開発されているが、情報検索用に使用する言語によって、おおよそ次のように分類される。

- ・COBOL、PL/1等を活用したもの(POL)
- ・独自のデータ・マネージメント言語(DML)を使うもの
- ・特殊なインプット様式を使うもの(FCS)
- ・POLは新しい言語を創出しないで、従来のコンパイラ言語を極力活用し、拡張したもので、COBOLとかPL/1で書かれたプログラムと共用して使うことができる。
- ・DMLは従来のコンパイラ言語と似た言語をデータ・マネージメント用に特に工夫、開発したもので、これによってプログラムが非常に簡略化される。
- ・FCSは企業や政府機関等でよく使われる申込書とか、仕様書のようにプログラムの様式をあらかじめ制定しておき、所定の欄にユーザーが書き入れれば良いように工夫したものである。

最も新しいデータ・マネージメント用ソフトウェアの代表的な例を表 3.2 に示す。

表 3.2 データマネジメント用ソフトウェア(例示)

名称	作成者	ファイルの種類	Pointer	host language	Direct access
TOTAL		master variable	calculating address→ master direct pointer variable chain two directional	any	可能
BSONP	IBM	master related chain	component chain where-use chain bi-directional	未完成	可能
MARS III	CDC	data base manages inquiry	Secondary index multiple field index	assembly ANSI cobol ANSI FORTRAN	
CZAR	Crown Zeller- back Co. (in- house)	Keyed file User structural file	index sequen- tional keyed file Chains user str. file	cobol	可能
DISK FORTE	Burroughs	Four Type randomized index-sequen- tional instructured	Three type a pointer link chain list	-	可能
SERIES	Information Systems Management Co.	unordered sequential hierarchical addressed	1st key 2nd key indexed	cobol	可能
IMS-2	IBM	hierarchiesl sequential hierarchical index sequen- tial " direct sequent. index direct	symbolic pointer direct address pointer	-	可能
DATA BASIC	Honeywell GE	small private file用 data definition を省略できる		-	可能

ところでアメリカの場合でもデータマネジメントシステムが実施面ではいまのところ発達しているとはいえないが、アメリカでは今後の発展のための基盤が、軍、宇宙産業、大学等のなかに、がっちりと建設されていることに注目すべきである。

日本は、オンラインシステムについても、データマネジメントシステムについてもまだほとんど初期の段階である。しかも日本では、'61年コボルを'65年になってはじめて使用するといった事実を考えれば、日本とアメリカにおけるソフトウェアに関する格差は大変なものといわなければならない。したがって、日本の場合はなによりもソフトウェアの開発に優秀なマンパワーが必要ということになる。

#### (4) コンピュータ技術の発展とデータ・バンク

いままで述べてきたように、コンピュータのシステムもオフ・ラインからオン・ラインへさらにリアルタイム処理、TSS処理へと発展してきた。その裏には真空管を用いた第1世代、トランジスタを用いた第2世代、ICを用いた第3世代へ、さらにIBM370に代表される第3.5世代へというハードウェアの発展とそれに伴うデータ・マネジメント用ソフトウェアの発展があったのは既述の通りである。

一方データ・バンクにおけるコンピュータの利用状況も併行して変わってきている。それはコンピュータの高速化とファイル容量の増大等が情報量の増大をいかに処理していくかという過程である。この項では、情報量の増大とその処理形態を述べ、データ・バンクの中核をなすデータの取扱いの推移を述べ、さらにデータ・バンク形成への問題点を述べる。

データ・バンクにおける情報処理システム別の一般的な情報処理量の推移と、それがどのような処理形態で処理されているかを示したものが図3.6である。

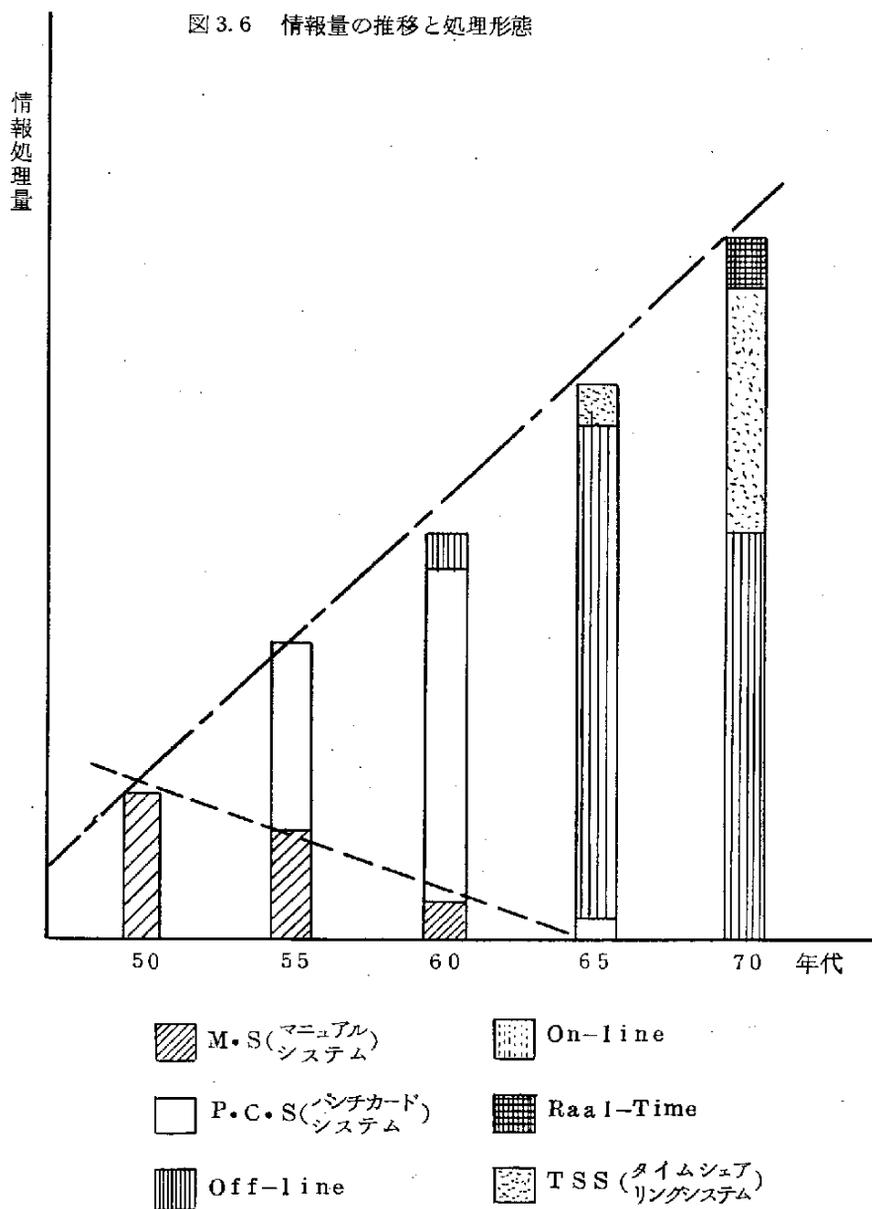
この図によれば、情報処理システムは、マニュアルシステム(M・S)からパンチカードシステム(P・C・S)へ、そしてEDPSへと移行してきた。EDPSになるとOff-line と On-line の二つのシステムに分かれていくようになるが、それはOn-line化へのみの移行だけでなく、Off-line の領域保存という形態でのOff-line から On-line への移行である。これは情報対象の特性に応じ処理システムが多様化していくことを示しており、なんでも全てがOn-line化の方向に処理システムが向かっているのではないことを語っているわけである。

アメリカでは、TSSによるリアル、タイム処理が、遠隔地の端末機からでも可能となったため、アメリカでのコンピュータ・ユーティリティに対するニーズは、多数のユーザーが利用し、かつ即時的な応答へと進んできていることも事実である。

この点が、わが国の大型コンピュータのバッチ処理を中心とする利用の仕方とのアメリカ

のそれとの差異が、両国の社会的ニーズの差異をもたらしているが、我が国においても回線の自由化がなされたので、上記の方向へと進んでゆくものと思われる。

この発展は、コンピュータによるデータの処理技術が高度化し、複雑化してきたことを示すものである。データ・バンクの中核をなすデータのコンピュータでの取扱いを考えると、次のような発展形態として捉えることができる。



データ・ファイル → データ・ベース → インフォメーション・マネージメント・システム (IMS)

現在、オンライン・システム時代に入っているが、おそらく、1970年代の前半が、データ・ベース、後半にIMSが実現されることになるであろう。

従来のデータ・バンクに使用されるファイルは、個々のアプリケーション業務（例、在庫管理、生産管理、販売管理など）ごとに別々に持たれていた。これではファイル内でデータの重複した部分や冗長な部分が多くなり、また更新する場合にも不都合を極めた。このような理由により、現在のデータバンク用ファイルは、統合化し、いくつかのアプリケーション適用業務プログラムが、それを共通に利用するような方向に進んでいる。即ち、プログラムのなかからデータの構造に関する部分を取り除き、独立したテーブル（データ記述テーブル：Data Description Table, Data Base Description etc.）にしようという方針がとられている。このような観点から考えると、『データ・ベースとは、複数のアプリケーション業務で使用される相互関係のあるデータの重複のない集合である』と定義できる。

一方、企業の情報ニーズを分析し、そのための情報を準備するという従来のアプローチには、限界があり、満足すべきシステムを作りあげることが、不可能であって、より積極的に“情報”そのものを管理し、いかなる要求にもダイナミックに応え得る新しいシステムの必要性が引きおこされてくるわけであるが、このシステムが、とりもなおさずIMS（インフォメーション・マネージメント・システム）のことである。このIMSにおいては、情報を収集し、配布するためのオンライン技術、タイムシェアリング技術などが不可欠のものであり、大量のデータを正確に蓄積し、迅速に加工するコンピュータの特性と人間のみの持つ創造性を融合させた人間機械系である。この意味で対話形言語の採用、プロッタやディスプレイなど、高度の端末機器の採用によって優れたマン・マシン・インターフェイスを確立することが重要である。

さて、最後にデータ・バンクの各構成要素に関する今後解決すべき技術上の問題点の主なものを列記すると次のようになる。

- (1) 高送：大容量かつ低価格の記憶装置が開発されること。
- (2) コンピュータ共同利用のソフトウェア技術を確立化すること（プログラムのモジュール化の徹底）
- (3) 端末装置として低価格で信頼性が高く、かつ操作性に優れたものを開発すること。
- (4) コンピュータを介して情報を加工するための簡単な（自然語に近い）言語が開発され

ること。(汎用言語の必要性)

(5) データ・バンクに関する秘密保護を完全にすること。

(6) わが国の場合には、さらに日本語(特に漢字)を効率的に扱う技術を開発すること。

以上、データ・バンク形成上解決すべき点を述べたが、データ・バンクを運営する上で、大量のデータを蓄積する能力をもったコンピュータ、データ・バンクのユーザとコンピュータを結ぶデータ通信設備、必要なデータを指示し検索するためのプログラムそれにユーザが手もとにおいて指示し結果を受け取るための端末器が必要である。これらの諸点について現在各方面で種々研究されており光メモリーの利用、効率的なモジュール編成の技法等徐々にではあるが解決の方向に向かっている。米国においてNASAや軍関係で開発されたこの方面の技術がデータ・バンク形式の技術に役立ったように、日本においてもこの方面の統一的な開発体制が望まれる。

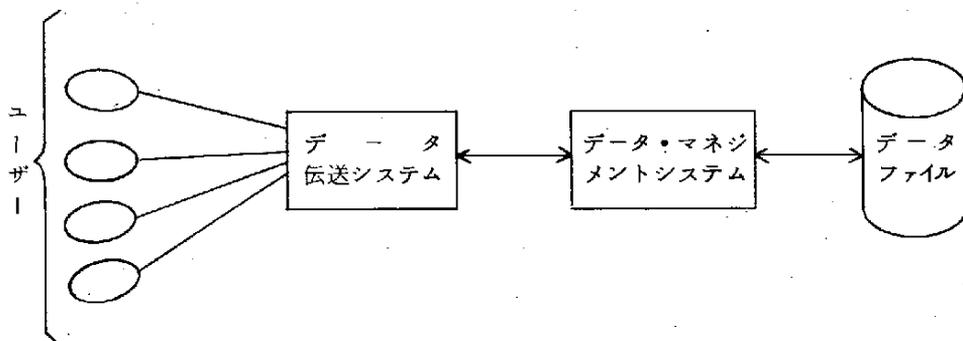
## Ⅱ 機能的側面からみたデータバンク

本節では1970年におけるアメリカのデータ・バンクの現状をクロス・セクション的に分析、評価することにする。

われわれがアメリカで調査したデータバンクは政府、銀行、証券会社、航空機会社、学校、研究所など非常に広範囲のもので、その内容や性格もそれぞれ異なっていて単純に比較分析することは難しい。また、われわれが調査したデータバンクがアメリカ全体として、一体どのような位置を占めるものなのか、また、将来これらのデータバンクは、どのような方向に向かって発展するのかを概観するのは容易でない。そこで、ここではデータバンクの母体に起因する差異をできるだけ捨象し、その機能的側面に重点をおいて大胆に単純化してみた。

まずデータバンクの基本的な構成要素を考えると、相互に重複がない、しかも更新、訂正等のデータ管理が行なわれている一つの総合データ・ファイルとそのユーザーの二つであろう。

つぎにこの両者を接続するものとして、データ伝送システムとデータ・マネジメントシステムが介在し、全体として下図のような構成をもって考えられる。

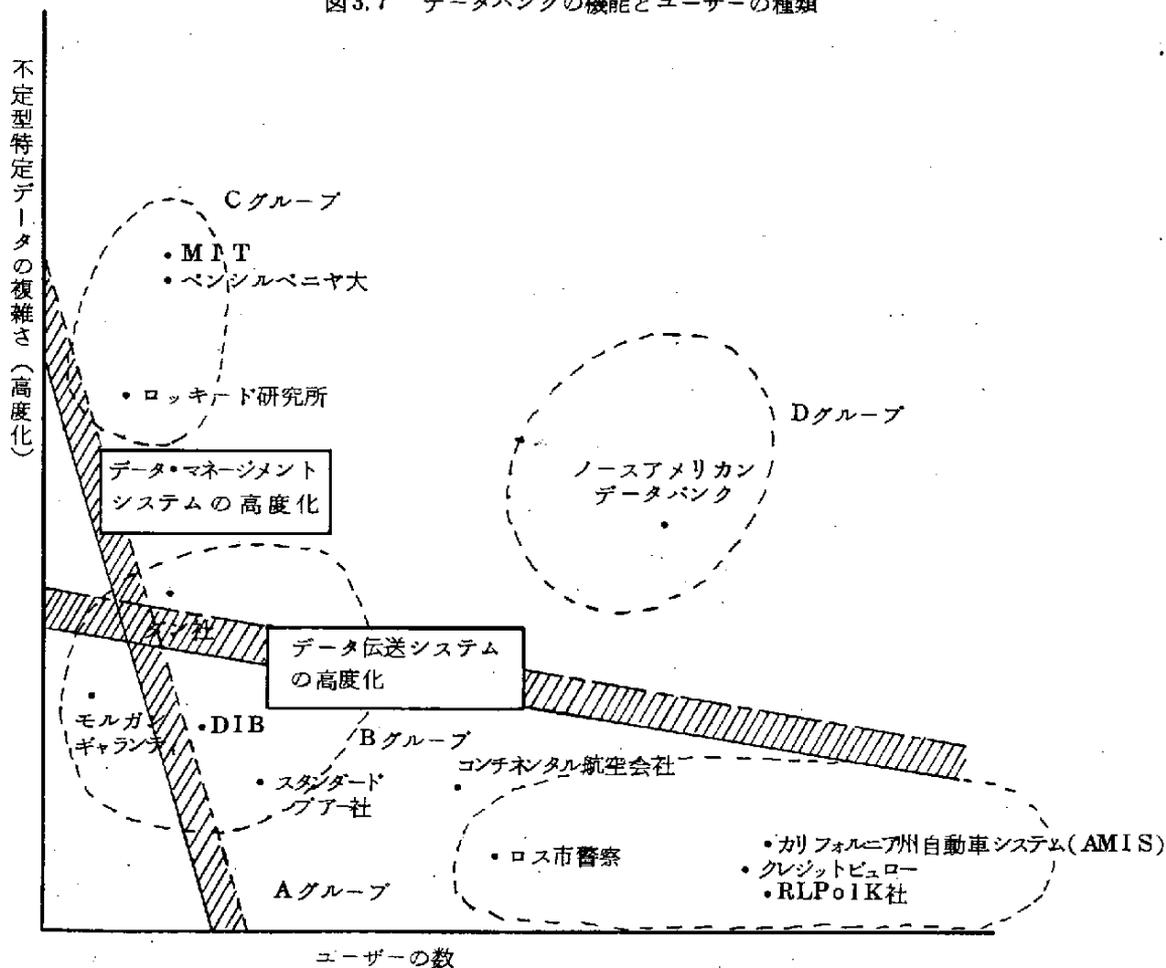


データ伝送システムは、端末機とデータ通信線を主体とするハード的な要素の強いものであるが、ユーザーの数が増大し、また遠隔地よりアクセスする要求が強まるにつれて、このデータ伝送システムの重要度は増大する。一方データ・マネジメントシステムは、データ収集、更新、訂正等のデータ管理や情報検索、作表、分析、秘密保護等を行なうもので、ソフト的な要素が強く、ユーザーの要求が高度化し、またそれに応ずるデータ・ファイルの構造が複雑化するにつれ、このシステムの重要性は高まってくる。

このようなデータ伝送システムとデータ・マネジメントシステムが、それぞれ機能的にみたデ

データバンクの量的拡大と質的向上を左右する要素であり、これらに関する技術進歩が、データバンクの発達の鍵を握っているといえよう。いま訪問したデータバンクを以上の観点から横軸にユーザーの数、縦軸にユーザーの要求度をとってプロットすると次の図のようになる。

図3.7 データバンクの機能とユーザーの種類



この図からデータバンクを位置づけてみると、A、B、C、Dの4グループに分けることができよう。これに対して、データ伝送システムおよびデータ・マネージメントシステムの技術進歩が、図のよな制約条件として作用している。

データ伝送システムの急速な発達により、Aグループに属するデータバンクが、次々と誕生している。一方データ・マネージメントシステムは、予想外に進歩が遅く本来ならばDグループにあるべきデータバンクがまだC、Bに低迷していた。

Dグループはデータ伝送システムとデータ・マネジメントシステムの両者を駆使した高度なデータベースであるが今回はノースアメリカン・カ所にすぎなかった。しかし、アメリカの大企業（製造業）、NASA、国防省関係には、この種のデータベースは決して少なくないと思われる。

日本の現状と比較してみると、データ伝送システムの未発達がネックとなって、現在Aグループでは国鉄、銀行等の例をのぞいて非常に遅れているが、データ伝送のネックさえ解消すれば、このグループのデータベースは早急にアメリカの水準に接近可能と思われる。しかし、C、Dになるとデータ・マネジメントシステムの技術格差が大きく、余程国家的な規模で人材と金を投入しないと、この格差を埋めることはむずかしい。

（註） グループA、B、C、Dは下記の性格を中心として分類したものである。

#### ◎グループA

ユーザーの数はきわめて多いが、データの内容は単純なグループで、カリフォルニア州政府の自動車登録（AMIS）やR. L. Polk 社などがその典型。

現在はアクセスに対して一度電話の中継という人間の要素を入れている場合が多く、技術的には自動化可能と思われるがサービス面でどちらが優れているか即断できない。

#### ◎グループB

政府統計のデータベース、連銀、スタンダード・プア社等

データがかなり専門的かつ一応印刷で発表されるデータが多く、直接データベースにアクセスを必要とするユーザーの数は限られている。しかしデータのupdatingや訂正がかなり面倒でしかも頻繁に行なり必要があり、現在、安全なデータベースはまだみられない。

（連銀の論文によればこの種のデータベース用ソフトウェアに対してコンピュータ・メーカーのサポートが少ないことも発達のネックになっているもよう。）

#### ◎グループC

各大学の図書館のIRシステム、ロッキード研究所のIRシステム。

データ・ファイルの内容が極めて複雑でかつユーザーの要求も非常に高度なため、現在そのソフトウェアの開発、研究が進められている段階。

実験的なものは完成（ロッキード研究所）しており、また理論的な研究が各大学で進められているが、一般的なソフトウェアパッケージは完成されていない。

#### ◎グループD

今回の訪問先ではノースアメリカン、ロックウェル社のみ。ただし、大企業、NASA、軍関係ではこの種のデータベースが多いと推測される。

データの数が多くかつユーザーの要求が高度なため、データ・マネジメントシステムは非常に複雑かつ高価なものになっている。一般公開されているものとしてはノースアメリカンのIMSがその代表例である。

このような分析のほかにも、次の考え方も成立し得る。

ユーザーの数とデータの複雑さの相関で捉えるならば

1. 組織内データベースのように、ユーザの数は少なくとも、データの内容が複雑であり、処理システムも複雑、高度であるようなデータバンクは現実には存在する。

そしてこれは、ファイルが、内部目的のために充分利用できる段階まで発展したとき、系列会社、関連業種をサービス範囲に収め、更に一般ユーザにも供給する傾向に移行する。

しかし、複雑さとは、当該時点におけるデータ処理技術、投下できるコストに限界があるため、ある限界がある。このため、図3.7の特性曲線Ⅰの周辺に事例が集まるであろう。

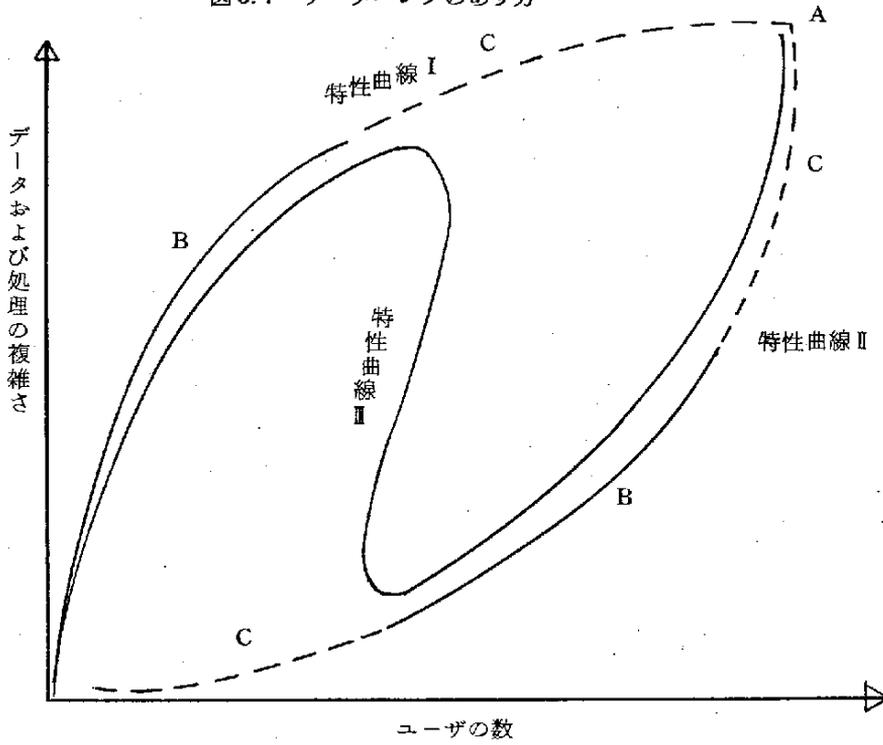
2. もう一つは、企業性データバンクのように、処理する技術が単純（ハードウェアやソフトウェアに要する経費が少なくすむため）であって、ユーザの数の多い所で成立するものがある。これはユーザの増加とある利潤増加によって歴史と共に資金供給をうけてより複雑なデータの処理や加工度の高いシステムの導入によって、更にユーザを獲得しようとする。このため、特性曲線Ⅱの周辺に事例が集まるであろう。

3. もし、この考え方が成立つならば、この2つの特性を、1つの解析関数で表わし得る。

それは特性関数値で近似できるであろう。

この考え方による分析を行なうことは、データが不足であって、各事例をプロジェクトする所までは行かなかったが、今後深く研究すべき課題の1つであろう。

図3.7 データバンクのあり方



- A: 或る時点で、技術コストの限界およびマーケットの限界から想定される理想点
- B: 現実に存在し得る部分
- C: 採算性とか、技術的とか、現実には成立し得ていない部分

### Ⅲ データバンクの成立パターン

データバンクは、その運営主体とその情報供給対象によって、各々のデータバンクの各機能のあり方は異なり、また、その成立過程にも大きなちがいがあある。

そこで、データバンクをその運営主体によって分類してみれば、次の3種類に分けることができる。

- 1) 公共的データバンク
- 2) 組織内データベース
- 3) 企業性データバンク

国家機関（行政、司法、立法の各部門に於ける）又は、地方公共団体がその中で発生する情報を一般に供給するために設置する場合、これを公共的データバンクと名づけることができよう。

この場合、データはその属する機関のオペレーショナル・ワークの中で多量に発生し、又特定のプロジェクトの遂行のために集められる。

このデータバンクの財政基盤は、その所属機関による資本投下と維持予算によって運営される。

この派生形態として、公共企業体、又は公共性の強い民間企業体が、その中で発生するデータを供給するために設置するものや、国又は地方公共団体が知的生産活動の振興、効率化のために設置するデータバンクがある。

後者の代表的な例としては、日本科学技術情報センター（JICST）などがある。このデータバンクの財政基盤は、社会資本としての性格が強いため、サービスコストを受益者が負担し、資本投下と維持費の一部は、公共的支出でなされている。

次に官公庁、研究機関、民間団体、民間企業体などが閉じたサークルの中だけでの利用のために設ける組織内データベースがある。

これは、たとえばある大学の研究室において学術的データを集め、同じテーマの研究者に配布したり、ある企業で系列会社に対してのみ情報を提供するような場合である。

営利を目的としたデータバンクを我々は企業性データバンクと名付ける。このデータバンクは、全く自由に経営が許され、自由競争と利用者の選択の中でみずからの方向を求める性格のものである。

以上の観点から日本とアメリカのデータバンクの現状を眺めてみると以下の表のとおりになる。

表 3.3 日本と米国のデータバンク機能（例示）

組織内データベース	企業性データバンク	公共的データバンク
<ul style="list-style-type: none"> <li>• ノースアメリカンロックウエル</li> <li>• ロッキードミサイル社</li> <li>• ニューヨーク連銀</li> <li>• ボストン小児科病院</li> <li>• ペンシルベニア大学</li> <li>• カリフォルニア州政府</li> </ul>	<ul style="list-style-type: none"> <li>• ダン &amp; ブラッドストリート社</li> <li>• マグローヒル社</li> <li>• スタンフォードL オプナー社</li> <li>• RLポーク社</li> <li>• スタンダード &amp; プアー社</li> <li>• アンシエテッドクレジットサービス</li> <li>• ウォートン経済予測研究所</li> </ul>	<ul style="list-style-type: none"> <li>• NLM</li> <li>• CAS</li> <li>• Foreign Commercial Service</li> <li>• NTIS</li> </ul>
<ul style="list-style-type: none"> <li>• 興 銀</li> <li>• 長 銀</li> <li>• 開 銀</li> <li>• トヨタ自販</li> <li>• 野村総研</li> <li>• 通産省</li> <li>• 警察庁</li> <li>• 労働省</li> <li>• 運輸省</li> </ul>	<ul style="list-style-type: none"> <li>• 交通公社</li> <li>• 日経データ 開発センター</li> <li>• 興 銀</li> <li>• 長 銀</li> <li>• 野村総研</li> </ul>	<ul style="list-style-type: none"> <li>• JICST</li> <li>• JETRO</li> <li>• アジア研究所</li> <li>• 特許情報センター</li> <li>• 住宅情報センター</li> </ul>

つきに、これらデータバンクがどのような生成過程を経て現在にいたっているかを考察してみよう。

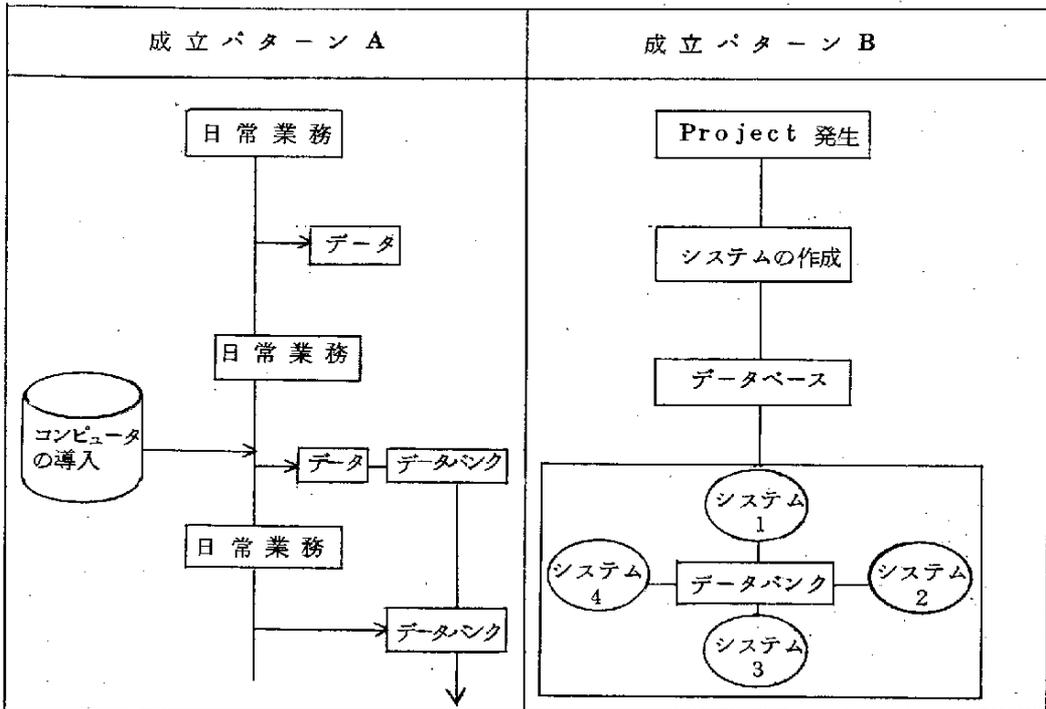
第2部に個々のデータバンクの事例が詳しく述べられているのであるが、これをみると、データバンクの成立過程は、大別すると次の2つに分けられる。

A：日常業務のプロセッシングによるデータ集積をもとにデータバンクに移行したもの。

B：高度な判断業務の要請による新しいデータベースの充実化により創られたデータバンク。

（JETRO、JICST、NASAのFAMEなど）

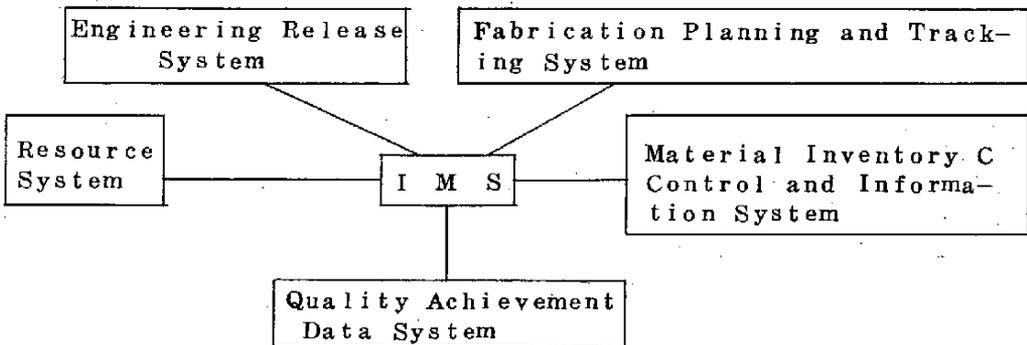
図 3.8 データバンクの成立パターン



成立パターンAは、アメリカの企業性データバンクがほとんどすべてこのパターンに該当し、日本でも開銀などがその代表的なケースである。

成立パターンBは、ノースアメリカンロックウェル社が、アポロのCSM(最上段の指令室)およびサターンS-IIのブースター製作に際し、開発した総合マネジメント・システム、これを運営するために情報処理を行うIMSと称するデータバンクがその最たるものである。

図 3.9 ノースアメリカンのMIS



その他では、カリフォルニア州政府の犯罪情報処理システム、自動車局の情報システムなどもあげられよう。

成立パターンAは企業性データベースの成立の自然過程とみなすことができるが、たとえその成立過程が同一であったとしても、その内容は例えばニューヨークのスタンダード・プア社はコンピュータスタットと呼ばれるデータベースであり、ダン&ブラッド・ストリート社はナンバリングシステム、マグローヒル、ダッチは建材データベースと特徴が実に鮮明である。

また、スタンダード・プア社やダッチのデータベースはデータを正確、迅速に提供するのが特徴であるが、ボストンのデータ・リソース社— この会社はTSS会社である — は経済統計データを企業に提供することのみならず、そのデータを使って経済予測などを行ないデータ処理がTSSにつながるようにサービスを提供している点、前者とは異った特徴を示している。

## IV パブリックセクターにおけるデータバンク

本節では、カリフォルニア州、西ドイツ、日本におけるパブリックセクターのデータバンクを、現状および構想に焦点をあて比較分析してみる。

まず、表、日米独のEDP化をご覧いただきたい。この表には日本、カリフォルニア州、西ドイツ連邦の3政府のEDP化の現状が数量的に示されている。

データ処理予算は、日本83億円（総予算の0.02%）、カリフォルニア州、299億円（総予算の1%）、西ドイツ連邦政府106億円（総予算 %である。

一般に、政府や個人企業の双方とも、データ処理の経費は、総予算のだいたい1%程度が適当であるといわれている。この言に従えば、カリフォルニア州がもっとも望ましい予算規模であるのに対して、日本、西ドイツはいまだしの感である。

コンピュータの設置台数、それに従事する人々の数などは、日本と西ドイツでは同じ規模である。カリフォルニア州の場合、予算規模の大きさに比して、コンピュータ設置台数が少ないのは、コンピュータの集中利用をはかっていることによるものである。

つきに、このようなコンピュータ利用の相違が、データバンクのあり方にどのような影響を与えているかを考察してみよう。

日本、カリフォルニア州、西ドイツ連邦の各々のパブリックセクターにおけるデータベースの現状は表3.5～表3.7に示されている。

これによれば、日本、カリフォルニア州、西ドイツ連邦の3政府間におけるデータバンクの開始の仕方が明らかに異なっていることがわかる。

一般にパブリックセクターにおけるデータバンクの成立手順を考えると、以下の3つのパターンが想定される。

表3.4 日、米、独のEDP化の実態

1970年

	日本政府	西ドイツ連邦政府	カリフォルニア州政府
億円 データ処理予算	83	106	299
セット コンピュータ設置台数	170	172	47
人数 従業員数	3400	3400	不詳
人 上級情報処理技術者	710	750	不詳

- 1) 固有業務（日本でいえば労働省の労働市場センター業務など）のEDP化よりデータバンクへ移行する型
  - 2) 内部管理業務のEDPよりデータバンクへ移行する型
  - 3) 1) と 2) の両者のパラレルなEDP化よりデータバンクへ移行する型
- これを図解すると以下のようになる。

図3.10

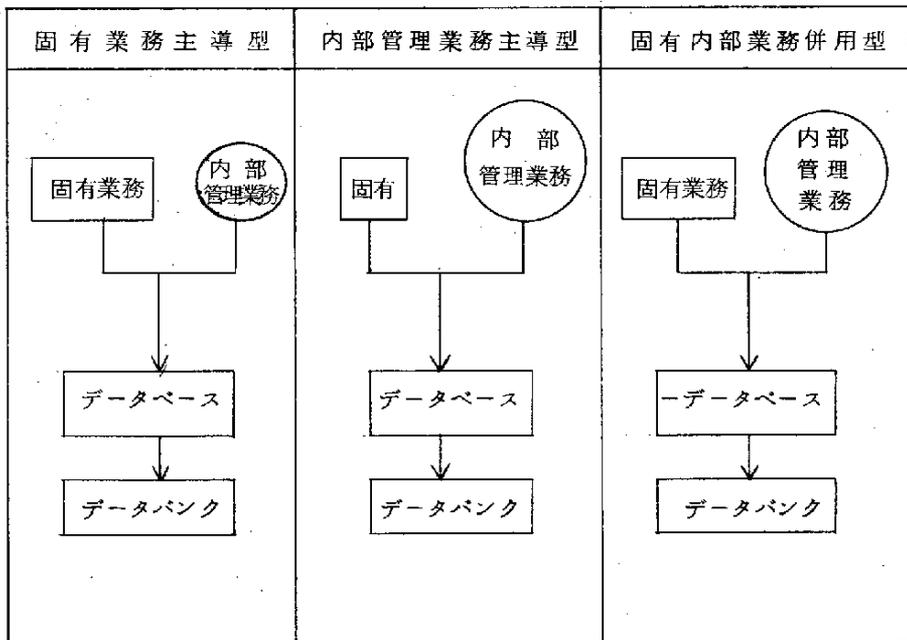


表 3.5 日本のパブリックセクターにおけるデータベース(例示)

関係官庁	概 要
警 察 庁	<ul style="list-style-type: none"> <li>◦ 運転免許不正取得システム</li> <li>◦ 犯例照会システム</li> <li>◦ 盗品照会システム</li> <li>On-line(部内用)</li> </ul>
経 済 企 画 庁	<ul style="list-style-type: none"> <li>◦ 経済データ総合管理システム</li> <li>バッチ計画中(部内用)</li> </ul>
外 務 省	<ul style="list-style-type: none"> <li>◦ 外交政策文章蓄積と Retrievalシステム</li> <li>バッチ(部内用)</li> </ul>
通 産 省	<ul style="list-style-type: none"> <li>◦ 一般経済情報蓄積と Retrievalシステム</li> <li>◦ 海外市場情報蓄積と Retrievalシステム</li> <li>◦ 産業経済統計蓄積と Retrievalシステム</li> <li>◦ 特殊文献一部蓄積と Retrievalシステム</li> <li>バッチ、一部 TSS(部内用)</li> <li>オフライン、ネットワーク(外部データ交換)</li> </ul>
運 輸 省	<ul style="list-style-type: none"> <li>◦ 許認可関係のコンピュータ化</li> <li>計画中(部内用)</li> </ul>
建 設 省	<ul style="list-style-type: none"> <li>◦ 国土開発に関する D・B</li> <li>計画中(部内用)</li> </ul>
大 蔵 省	<ul style="list-style-type: none"> <li>◦ 関税統計販売</li> <li>バッチ(関税協会を通じて外販売)</li> </ul>
労 働 省	<ul style="list-style-type: none"> <li>◦ 職業紹介と労災のレジストレーション</li> <li>◦ 指定統計の改正への動き</li> <li>(On-line)部内用</li> </ul>
法 務 省	<ul style="list-style-type: none"> <li>◦ 登記事務コンピュータ化</li> <li>バッチ計画中(部内用)</li> </ul>
自 治 省	<ul style="list-style-type: none"> <li>◦ 地方自治情報システム</li> <li>◦ 税務関係等、バッチ計画中(部内)</li> </ul>
運 輸 省	<ul style="list-style-type: none"> <li>◦ 自動車登録システム</li> <li>◦ 自治省、警察庁等と関連バッチ、オフラインデータ交換システム</li> </ul>
科 学 技 術 庁	<ul style="list-style-type: none"> <li>◦ 防災センターの D・B</li> <li>計画中</li> </ul>
国 会 図 書 館	<ul style="list-style-type: none"> <li>◦ 米国議会図書館の MARK II を導入した図書館録</li> <li>◦ 書誌的事項の整理、サービス</li> </ul>
総 理 府 統 計 局	<ul style="list-style-type: none"> <li>◦ 統計 D・B</li> <li>◦ 対外サービスを計画</li> </ul>
日 銀	<ul style="list-style-type: none"> <li>◦ 統計データ、D・B</li> <li>バッチ(部内用)</li> </ul>
ア ジ 研	<ul style="list-style-type: none"> <li>◦ 海外貿易統計</li> <li>◦ 国連と O・E・C・D と NETワーク</li> </ul>
J E T R O	<ul style="list-style-type: none"> <li>◦ 海外経済情報 D・B</li> <li>◦ 貿易あっせん+海外市場情報、民間サービス、バッチ</li> </ul>
特許情報センター	<ul style="list-style-type: none"> <li>◦ 特許情報の民間サービス</li> <li>◦ 海外とのネットワーク作業(開発中)</li> </ul>
農 林 省	<ul style="list-style-type: none"> <li>◦ 生鮮食料品情報システム</li> <li>◦ 物価コントロールシステム(開発中)</li> </ul>
住宅産業情報センター	<ul style="list-style-type: none"> <li>◦ 建設省、通産省、関係産業界等が関連、民間サービス用(開発中)</li> </ul>
J I C S T	<ul style="list-style-type: none"> <li>◦ 科学技術情報の収集、抄録の電算機編集、IRサービス、MTサービス</li> </ul>

表3.6 カリフォルニア州政府のデータベース(例示)

管 轄	シ ス テ ム 名	概 要
農 業 部		農業害虫駆除
教 育 部	CEIMS (カリフォルニア教育) 情報管理システム	教育管理のためデータベースを 与える
財 政 局		歳入と人口の推計
健 康 管 理 局		健康管理サービスの資格審査
ハ イ ウ ェ イ パ ト ロ ー ル 局	AUTO - STATIS	盗難車輛の情報蓄積
人 間 関 係 局		コンピュータによる社会情報シス テムの全国規模のデモンストレーション
人 材 開 発 部	WIN (自動職業奨励システム)	職業奨励
司 法 局	カリフォルニア法律施行通信システム	
精 神 衛 生 局		州病院患者データバンク
州 兵 局		軍人データ・ファイル
自 動 車 局		
公園とレクリエーション局	公園とレクリエーション情報システム	
公 共 事 業	交通事故と監視分析システム (TASAS)	
社 会 福 祉 局	州教員退職システム	
水 資 源 局	基礎データシステム	
査 定 平 準 局		資産税査定標準
大気資源委員会		大気状態データのコンピュータに よる処理
会 計 検 査	地方政府会計事務プログラム	
公 衆 衛 生 局		職業病報告

表 3.7 西ドイツのパブリックセクターのデータバンク(計画)

<p>新聞情報局の領域 データ読取機械 押しボタン及び瞬間印字機から成る 端末装置</p>	<p>1970年以降、「政治的内容をもつ情報の記憶と再生」のデモン ストレーション・データ処理計画を実施しており、現在は、データ処 理及びマイクロフィルムから成っている統一的情報供給システム(I BS)の設立作業をおこなっている。</p>
<p>連邦司法省の領域 ミュンヘンのドイ ツ特許庁における 作業行程の合理化</p>	<p>特許制度情報システムの作成          手続き情報 { 内的情報                            { 外的情報          技術情報 { 図書試験材料の開発の諸方法                            { 本来の特許書類          (連邦中央登録局の設置)          前科表 - 電子工学的データ処理の導入により中央登録局に                    統合する          (法律データバンク)          立法、判決、法学文献の分野における無数の情報を適宜整理して、          これを実用に適するものにするようオートメーション化された情報          システムを開発する。</p>
<p>連邦交通省の領域 資料整備における 情報再生手続き</p>	<p>(連邦道路局)          交通の円滑さと安全性の向上のために策定すべき諸原則は、技術          的、科学的計算とともに、電子工学的データ処理の助けによる計算          と測定を活用と資料整備における近代の情報再生手続きの導入が必          要となる。          (交通データバンク)          交通政策に必要な重要な資料を記憶する。</p>
<p>連邦労働社会省の 領域</p>	<p>保険技術面で有用なデータはすべて機械的に把握され、これをす          べての該当管理体に提供する。</p>
<p>連邦内務省の領域 データバンク</p>	<p>(連邦統計局)          統計データバンクの設立          ・科学的分析          ・モデル計算          ・租税及び財政政策          ・農業          ・交通制度          ・都市計画          上記の分野に予測と計画作成の為の援助を与える。</p>

この分類によれば、日本政府のデータバンク生成のプロセスは、固有業務主導型で、西ドイツ連邦政府は内部管理主導型である。両者の併行推進型がカリフォルニア州政府のデータバンクのそれである。

内部管理業務のEDP化よりデータバンクへ移行する型は、EDP化の進展は顕著であるが、データベースの充実がむずかしく、したがってデータバンクへの移行が困難である。

固有業務主導型は、データベースの充実がはかりやすいが、内部管理業務とのEDP化が立遅れ気味となり、この点からの影響がデータバンクへの成立の阻害要因となる。

カリフォルニア州政府のデータバンクの成立手順は、データ処理予算が日本の3倍もの約300億円費いやしているということもあって内部管理業務も固有業務もともにEDP化の促進がなされ、同時に両者のデータベース相互間の関係も考慮されている。それ故、パブリックセクターにおけるデータバンクの成立手順としては地味で手がたいがもっとも望ましい型で進められているといえよう。

表3.7の西ドイツ連邦政府のデータバンクは「電子計算機利用に関する技術研究会資料」にもとづいている。

# む す び

## わが国におけるデータバンクの発展と展望

### 1. 日米の現状比較

だれでもが、何時、何処からでも、必要な情報を入手できるという人類の夢に、わが国のデータバンクも答え得るようになれるものであろうか。

まず、アメリカについてみよう。第2部で数多くの事例を示したように、アメリカでは現在までにデータバンクが誕生しており、さらに、さまざまなタイプのデータバンクが開発の途上にある。

とくに、組織の大小を問わず、いたる所に端末機が置かれ、簡単に遠隔地からアクセスしたり結果をディスプレイに映し出している光景は、アメリカ調査団の一行に強烈な印象を与えた。

アメリカにおいて現在完成し、また開発中の大部分のデータバンクは、従来のデータ処理機能の進歩が自然に結実して行ったものである。しかしそのほかに新しい特定のコースの発生に応じて開発されたデータバンクも、まだ数は少ないが、いくつかあるアポロ製作用に開発されたノースアメリカン社のデータベースや、センサデータの多角的活用をめざして開発中の Geographic Base File 等がその例である。

また、アメリカでは、軍事防衛システムや社会開発という背景がデータバンクの画期的な発展を促進しているのは注目しなければならない。そこでは豊富なコンピュータ技術と通信技術を活用し、データバンクの開発に、龐大な資金とマンパワーを集中している。

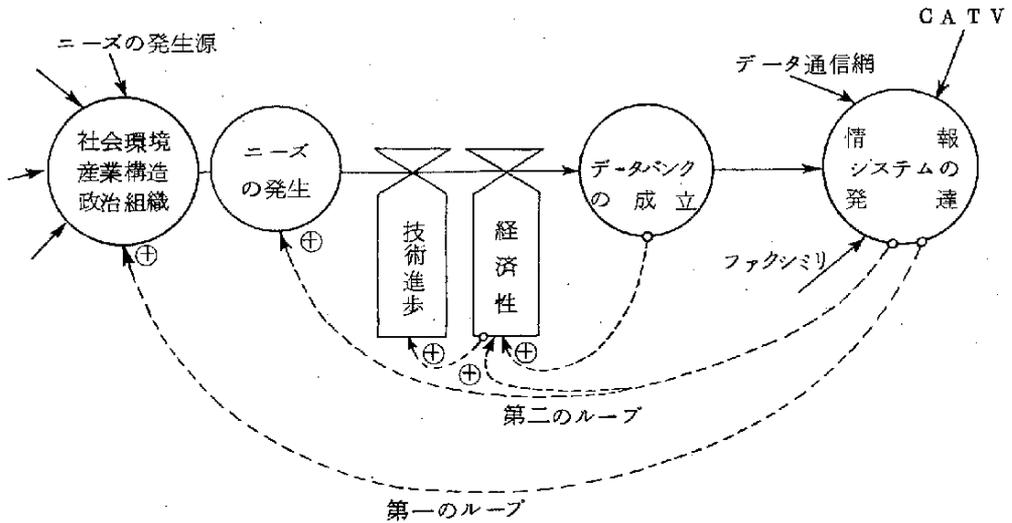
このようにアメリカは、「だれでもが、何処からでも」という理想に向って、いろいろな方向から確実に歩を踏み出しているといえよう。

このアメリカと比較すると、わが国のデータバンクは平均して5～6年ほど遅れているといわれる。現在公共的データバンクの僅かな事例を除いては、本格的なデータバンクと呼べるものは存在していない。しかし、第2部で事例を示したように、その萌芽は多方面に渡って見ることができるので、今後の発展可能性を蔵しているとはいえる。しかしながら現段階での日米格差は、外見上みられる遅れよりも、その技術的蓄積や情報処理システム全体の背景等を考慮に入れると実質的な差はより大きいかもしれない。

### 2. データ・バンクの発展メカニズム

わが国のデータバンクの今後の発展過程を考察するためには、社会の進展・変化のなかに、データバンクを位置づけてみる必要がある。(第3.1.1図参照)。この図において注すべきことは、次のような大きなフィード・バックの存在することである。

第3.11図 データバンクの成長メカニズム



(i) まず政治・経済・社会のあらゆる活動を通して情報が発生するとともに、この情報の選択流通速度の向上、等を目的として、データバンクに対するニーズが発生する。

(ii) このニーズは、技術的、経済的、あるいは 度的制約条件のもとで、徐々にデータバンクの成立に結びつく。

(iii) そのデータバンクが、再び政治・経済・社会の活動力を高める方向にフィードバックされていくことを示したものである。

このような大きなフィードバックループのほかに、一つのデータバンクの成立が、次のデータバンクの成立を促がし、ネットワークの網を拡大しながら、細胞分裂のように急速に、増殖発展するメカニズムも含まれている。これら2つの作用が相まって、データバンクがダイナミックに発展して行くものと考えられる。

### 3. わが国におけるデータバンク発展の基盤

そこで、この発展のメカニズムが、わが国の実情において、いかに展開されるかを見よう。

まず、ニーズの発生源である政治・経済・社会の活動力という点からみれば、わが国は欧米諸国に例を見ないポテンシャルを蔵している。1億の人口が同一民族であり、同じ言葉を話し均質な文化を有し、高い教育レベルを維持している。朝日、毎日等の新聞発行部数約500万部ということは、ニューヨークタイムズの90万部、ロンドンタイムズの40万部と比較して、日本文化の均質性と、情報の拡がりを象徴しているといえよう。

G.N.Pは現在すでに2000億ドルで、アメリカ、ソ連について世界第3位であり、1985年には現在のような成長を継続すると、一人あたりGNPが14,000ドルと、世界第1位になるといわれる。データバンクの本格的な展開が10年後と考えれば、わが国のポテンシャルな活動力は、極めて重要な意味を持つことになる。

また、わが国の国土は狭く、天然資源にも乏しい。1平方Kmあたり人口密度は270人で、アメリカの21人に比べると約10倍にあたり、もし可住面積あたり人口密度にすれば、その差は30～40倍にも拡大しよう。これを1平方Kmあたり所得で比較すれば、現在すでに世界最高であり、今後の成長を考えれば驚くべき高密度の未踏社会に向って進んでいることは明らかである。

しかし、現段階ではアメリカにくらべて、データ流通は少い。これはアメリカが異民族の複合社会であること、それにアメリカでは情報価値の観念が進んでいて、情報交換の経済性が確立していることが大きな原因である。わが国のような高密度社会では、質の高いデータが今後要求されることになろう。そこに、単なるデータ流通とは違ったデータバンクの機能を必要とすることになる。

つぎに、データバンク成立の原動力である技術進歩の面からみると、データバンクの技術ではわが国は遅れているが、その総合的な技術基盤では、アメリカとの差は非常に接近しており、将来のポテンシャルからみて、むしろわが国の方が強い面も少くない。

このように、わが国におけるデータバンク発展の基盤としては、(1)経済力およびその将来の発展予想、(2)情報化指標の普及、(3)技術力の格付、などはプラスの面であり、情報価値観の確立していないことはマイナスの面といえる。この後者の点はわが国のホモジニアスな文化と高密度な社会に起因することが大きい。今後は当然、この価値観も変更して行くものと考えられる。

#### 4. わが国におけるデータバンク発展のための起動力

わが国のデータバンクは、現在まだ萌芽状態である。しかしアメリカにおけるデータバンクの発展からみても、おそらく5～6年後にはデータニーズの多面的発生に伴い、情報システムの発達により、社会環境のなかにデータバンクが根を下すものと推察される。また、その方向に発展の起動力がどこにあるかを格付してみたい。

第1は、民間データバンク発展への期待である。アメリカでは現在P.L.ボーク社やマグローヒル社等にみられるように、企業性データバンクが活動しているが、わが国ではまだ日本経済データ開発センターや日本不動産取引情動センター等が誕生したばかりである。これはデータニーズの歴史的基盤のちがひによる点が多いが、わが国の企業の持つバイタリティや優れた適応力から考えて、自由競争条件の下で、多様な方面に開花し、発展していく能力を十分に備えて

いると思われる。しかも、わが国のような社会の発展・変化が著しい所で、ソシアル・ニーズに  
適応できるのは、企業性データバンクであるはずだ。多くの資本、すぐれた技術それに長年月の  
経験を投じて、権威ある民間データバンクの発展が望まれる。

第2は、公共的データバンクの重要性とその責任である。政府、地方公共団体の業務を通じる  
甚大なデータの発生は、適切な整理と流通チャネルを持たせることによって、国民の貴重な公共  
財とすることができるし、公共団体はそうする責任があると云えよう。

アメリカの連邦政府にはナショナル・データバンクの構想があり、西ドイツ政府にはデモテ  
レーション・プログラムに基づくデータバンクの計画がある。しかし、これらの理想的システムの  
計画からの発想は、データバンクの実現に必ずしも成功しているとはみられない。政府のデー  
タバンクの具体的な開発方法としては、カリフォルニア州の経験が、貴重な教訓になると思われる。

データバンクは、その性格から考えて、周到な設計と計画性が必要なように見える。とくに、  
公共的データバンクでは、将来構想に基き、すべての公共データを統括して、もっとも効率の高  
いシステムの設計に進むべきだと考えがちである。

カリフォルニア州政府で、実際に開発に従事したマクドナルド氏によれば、自動車情報システ  
ム（AMIS）の例のように理論の飛躍をさけ、現実のニーズに応じて、確実に保証された技術  
を採用しながら、一步一步着実に開発を進めている。現実問題としては各部門の権限や、組織変  
更に対する抵抗等、技術以前の問題が少くない。また技術そのものも現在また確実なものも少く  
とくにデータマネージメント用ソフトウェアには解決すべき多くの問題があり、彼等は地道な  
研究を続けている。

技術進歩が激しく、社会が要求するニーズの変化も決して無視できない条件の下では、デー  
タバンクの開発に綿密な調査・計画とともに、理想論に走らず、できるだけシステム・フレキシビ  
リティを持たせながら現実的な解決策を積みあげていくほか道はない。

一見遅々たるように見えながら、こうして開発されてきたカリフォルニア州政府のデータバン  
クは、わが国の公共的データバンクの開発に大に参考になると思われる。

第3は、データバンクに対する新しいニーズの発生である。わが国は現在、経済成長の盾の裏  
側として、都市再開発や公害発生等幾多の問題に直面している。これらを長期的に組織的に解決  
するためには、まず情報システムを確立する必要がある。その武器として、データバンクは欠く  
ことができない。従ってここからデータバンク発展の起動力がうまれる。

カリフォルニア州では、交通情報システム、犯罪防止システム、あるいは廃棄物管理システ  
ムを、1965年に着手し、それらのシステムを実践するために、自動車情報（AMIS）や犯罪  
情報（CLETS）等のデータバンクを着実に開発し、1970年には稼動するに到っている。

わが国でも、国家的なプロジェクトから発生する公共的データバンク・システムの開発が期待されるところである。

第4は、データバンクを接続するネットワークの強化である。データバンクの価値は、そこに含まれるデータを正しく認識し、評価し、加工分析し、要求に応じて結果を伝達する能力によって定まり、当然のことながら、分化専門化の必要が高まってくる。一方多様な要求に答えるためには、その要求に見合ったデータバンクを探し出さなくてはならない。このために高度な選択機能を持ったネットワークの必要性が浮び上ってくる。またこのネットワークは、情報システム全体のフレキシビリティを保つためにも重要な要素と考えられる。

データバンクのネットワークとは機能的に接続することであって、必ずしも通信回線による接続を意味しない。しかし、通信回線が近く民間に解放されることは、データバンクのネットワークを促進することも確かである。われわれは巨大な、総合的なデータバンクの開発をめざすのではなく、多様な、専門的なデータバンク・ネットワーク開発に取組まなければならない。

第5は、情報化政策の確立による処理である。データバンクのネットワークが確立されるときは、それは社会の神経網となるものである。そのような役割を果し得るようにするとともに、その暴走を防止する政策をあらかじめ確立しておく必要がある。

そのために、技術やフォーマットの標準化・標準コードの普及、あるいはデータ・マネジメントなどのソフトウェア技術の保護などについて政策を確立しなければならない。アメリカでは情報化の進展につれて、プライバシー保護が大きな問題となっている。わが国においても、情報は今後オープンにされるべきだが、そのなかにプライバシーの問題はどうなるのか、事前に検討しておくべきであろう。

このように一面において情報化を促進し、他面において弊害を防止するために、新しい法制化が必要になろう。また、発想を変えて、上記の諸問題を包括する「情報憲章」とも呼ぶべき基本的な職業倫理を成文化することも検討に値いすることであろう。

#### 5. データバンク発展のための提言

以上のような考察をもとにわが国におけるデータバンクの発展のためには、特に提言したいと思うことは、次の五綱目である。

第1に民間データバンクの自発的な発展をさまたげず、これを助長する方向の政策がのぞましい。ここでは当然競争原理による発展が主役を演じる。

第2に公共データバンクの発展のためには、国家的な施策が要望される。ここでは理想にとらわれず地道な努力を早急かつ着実にすすめるべきである。

第3にデータバンクに対する新しいニーズの発生に伴い情報システムの確立が肝要となる現状

に即して、データバンク形成に 的の努力をすべきである。

第4にデータバンクを接続する高度の選択機能をもつネット・ワークの強化に力をそそぐべきである。

第5にこれら全体に対する施策の基盤として、情報化政策の確立とその実施をはかることが特に要望される。

